PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/55, 9/18, A23D 7/00, A23L 1/10

A1

(11) International Publication Number: WO 00/28044

(43) International Publication Date: 18 May 2000 (18.05.00)

(21) International Application Number: PCT/US99/26789

(22) International Filing Date: 10 November 1999 (10.11.99)

(30) Priority Data:
09/189,486 10 November 1998 (10.11.98) US

(71) Applicants: NOVO NORDISK BIOTECH, INC. [US/US]; 1445 Drew Avenue, Davis, CA 95616 (US). NOVO NORDISK BIOINDUSTRY, LTD. [JP/JP]; Makuharui Techno Garden CB-6, 3, Nakasa, 1-chome, Chiba-shi 261-01 (JP).

(72) Inventors: BERKA, Randy, M.; 3609 Modoc, Davis, CA 95616 (US). REY, Michael, W.; 605 Robin Place, Davis, CA 95616 (US). BYUN, Tony; Apartment B-1, 2680 Sycamore Lane, Davis, CA 95616 (US). ITAMI, Ryoko; 6-1, Wadamichi, Sukagawa-shi, Fukushima-ken 962-0818 (JP). TSUTSUMI, Noriko; 3-2-16-2, Higashi Sugano, Ichikawa-shi, Chiba-ken 272-0823 (JP). KLOTZ, Alan; 725 Collier Drive, Dixon, CA 95620 (US).

(74) Agents: ZELSON, Steve, T. et al.; Novo Nordisk of North America, Inc., 405 Lexington Avenue, Suite 6400, New York, NY 10174 (US).

(81) Designated States: AE, AL, AU, BA, BB, BG, BR, CA, CN, CU, CZ, EE, GD, GE, HR, HU, ID, IL, IN, IS, JP, KP, KR, LC, LK, LR, LT, LV, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, SL, TR, TT, UA, UZ, VN, YU, ZA, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: POLYPEPTIDES HAVING LYSOPHOSPHOLIPASE ACTIVITY AND NUCLEIC ACIDS ENCODING SAME

(57) Abstract

The present invention relates to isolated polypeptides having lysophospholipase activity and isolated nucleic acid sequences encoding the polypeptides. The invention also relates to nucleic acid constructs, vectors, and host cells comprising the nucleic acid sequences as well as methods for producing and using the polypeptides.

POLYPEPTIDES HAVING LYSOPHOSPHOLIPASE ACTIVITY AND NUCLEIC ACIDS ENCODING SAME

Background of the Invention

Field of the Invention

The present invention relates to isolated polypeptides having lysophospholipase activity and isolated nucleic acid sequences encoding the polypeptides. The invention also relates to nucleic acid constructs, vectors, and host cells comprising the nucleic acid sequences as well as methods for producing and using the polypeptides.

Description of the Related Art

Phospholipases are enzymes that participate in the hydrolysis of phospholipids which consist of a glycerol backbone with two fatty acids in an outer (sn-1) and the middle (sn-2) positions and esterified with phosphoric acid in the third position. The phosphoric acid may, in turn, be esterified to an amino alcohol.

Several types of phospholipase activity can be distinguished which hydrolyze the fatty acyl moieties. Phospholipase A1 and A2 catalyze the deacylation of one fatty acyl group in the sn-1 and sn-2 positions, respectively, from a diacylglycerophospholipid to produce lysophospholipid. Lysophospholipase (also called phospholipase B by the Nomenclature Committee of the International Union of Biochemistry on the Nomenclature and Classification of Enzymes {Enzyme Nomenclature, Academic Press, New York, 1992}) catalyzes the hydrolysis of the remaining fatty acyl group in a lysophospholipid. A phospholipase B has been reported from *Penicillium notatum* (Saito *et al.*, 1991, *Methods in Enzymology* 197:446-456) which catalyzes the deacylation of both fatty acids from a diacylglycerophospholipid and intrinsicly possesses lysophospholipase activity.

Fungal enzymes with phospholipase activity have been reported from various sources, including *Cryptococcus neoformans* (Chen *et al.*, *Infection and Immunity* 65: 405-411), *Fusobacterium necrophorum* (Fifis *et al.*, 1996, *Veterinary Microbiology* 49: 219-233), *Penicillium notatum* (also known as *Penicillium chrysogenum*; Kawasaki,

nucleotides 214 to 2061 of SEQ ID NO. 1 or nucleotides 49 to 1944 of SEQ ID NO. 15; (iii) a subsequence of (i) or (ii) of at least 100 nucleotides; or (iv) a complementary strand of (i), (ii), or (iii);

(c) an allelic variant of (a) or (b); and

(d) a fragment of (a), (b) or (c), wherein the fragment has lysophospholipase activity.

The present invention also relates to isolated nucleic acid sequences encoding the polypeptides and to nucleic acid constructs, vectors, and host cells comprising the nucleic acid sequences as well as methods for producing and using the polypeptides.

Brief Description of the Figures

Figures 1A, 1B, and 1C show the cDNA sequence and the deduced amino acid sequence of a *Fusarium venenatum* lysophospholipase (SEQ ID NOS. 1 and 2, respectively).

Figure 2 shows a restriction map of pDM181.

Figure 3 shows a restriction map of pSheB1.

Figure 4 shows a restriction map of pRaMB54.

Figure 5 shows the pH activity profile of a *Fusarium venenatum* lysophospholipase.

Figures 6A, 6B, and 6C show the cDNA sequence and the deduced amino acid sequence of a *Fusarium verticillioides* lysophospholipase (SEQ ID NOS. 15 and 16, respectively).

Detailed Description of the Invention

Polypeptides Having Lysophospholipase Activity

The term "lysophospholipase activity" is defined herein as a carboxylic ester hydrolysis activity which catalyzes the deacylation of one or both of the fatty acyl groups in the sn-1 and sn-2 positions of a diacylglycerophospholipid. For purposes of the present invention, lysophospholipase activity is determined by incubating the lysophospholipase with lysolecithin (or L-α-lysophosphatidylcholine) in the presence of

lysophospholipase activity. In another preferred embodiment, the polypeptide of the present invention consists of the amino acid sequence of SEQ ID NO. 2. In another preferred embodiment, the polypeptide consists of amino acids 38 to 654 of SEQ ID NO. 2 or an allelic variant thereof; or a fragment thereof, wherein the fragment has lysophospholipase activity. In another preferred embodiment, the polypeptide consists of amino acids 38 to 654 of SEQ ID NO. 2.

A fragment of SEQ ID NO. 2 is a polypeptide having one or more amino acids deleted from the amino and/or carboxyl terminus of this amino acid sequence. Preferably, a fragment contains at least 500 amino acid residues, more preferably at least 550 amino acid residues, and most preferably at least 600 amino acid residues.

Preferably, the polypeptides of the present invention comprise the amino acid sequence of SEQ ID NO. 16 or an allelic variant thereof; or a fragment thereof that has lysophospholipase activity. In a more preferred embodiment, the polypeptide of the present invention comprise the amino acid sequence of SEQ ID NO. 16. In another preferred embodiment, the polypeptide of the present invention comprises amino acids 17 to 648 of SEQ ID NO. 16, which is the mature polypeptide of SEQ ID NO. 16, or an allelic variant thereof; or a fragment thereof, wherein the fragment has lysophospholipase activity. In another preferred embodiment, the polypeptide of the present invention comprises amino acids 17 to 648 of SEQ ID NO. 16. In another preferred embodiment, the polypeptide of the present invention consists of the amino acid sequence of SEQ ID NO. 16 or an allelic variant thereof; or a fragment thereof, wherein the fragment has lysophospholipase activity. In another preferred embodiment, the polypeptide of the present invention consists of the amino acid sequence of SEQ ID NO. 16. In another preferred embodiment, the polypeptide consists of amino acids 17 to 648 of SEQ ID NO. 16 or an allelic variant thereof; or a fragment thereof, wherein the fragment has lysophospholipase activity. In another preferred embodiment, the polypeptide consists of amino acids 17 to 648 of SEQ ID NO. 16.

A fragment of SEQ ID NO. 16 is a polypeptide having one or more amino acids deleted from the amino and/or carboxyl terminus of this amino acid sequence. Preferably, a fragment contains at least 500 amino acid residues, more preferably at least 550 amino acid residues, and most preferably at least 600 amino acid residues.

An allelic variant denotes any of two or more alternative forms of a gene

high stringency conditions with a nucleic acid probe which hybridizes under the same conditions with (i) nucleotides 214 to 2061 of SEQ ID NO. 1 or nucleotides 49 to 1944 of SEQ ID NO. 15, (ii) the genomic DNA sequence containing nucleotides 214 to 2061 of SEQ ID NO. 1 or nucleotides 49 to 1944 of SEQ ID NO. 15, (iii) a subsequence of (i) or (ii), or (iv) a complementary strand of (i), (ii), or (iii) (J. Sambrook, E.F. Fritsch, and T. Maniatus, 1989, *Molecular Cloning, A Laboratory Manual*, 2d edition, Cold Spring Harbor, New York). The subsequence of SEQ ID NO. 1 or SEQ ID NO. 15 may be at least 100 nucleotides or preferably at least 200 nucleotides. Moreover, the subsequence may encode a polypeptide fragment which has lysophospholipase activity. The polypeptides may also be allelic variants or fragments of the polypeptides that have lysophospholipase activity.

The nucleic acid sequence of SEQ ID NO. 1 or SEQ ID NO. 15 or a subsequence thereof, as well as the amino acid sequence of SEQ ID NO. 2 or SEQ ID NO. 16, or a fragment thereof, may be used to design a nucleic acid probe to identify and clone DNA encoding polypeptides having lysophospholipase activity from strains of different genera or species according to methods well known in the art. In particular, such probes can be used for hybridization with the genomic or cDNA of the genus or species of interest, following standard Southern blotting procedures, in order to identify and isolate the corresponding gene therein. Such probes can be considerably shorter than the entire sequence, but should be at least 15, preferably at least 25, and more preferably at least 35 nucleotides in length. Longer probes can also be used. Both DNA and RNA probes can be used. The probes are typically labeled for detecting the corresponding gene (for example, with ^{32}P, ^{3}H, ^{35}S, biotin, or avidin). Such probes are encompassed by the present invention.

Thus, a genomic DNA or cDNA library prepared from such other organisms may be screened for DNA which hybridizes with the probes described above and which encodes a polypeptide having lysophospholipase activity. Genomic or other DNA from such other organisms may be separated by agarose or polyacrylamide gel electrophoresis, or other separation techniques. DNA from the libraries or the separated DNA may be transferred to and immobilized on nitrocellulose or other suitable carrier material. In order to identify a clone or DNA which is homologous with SEQ ID NO. 1 or SEQ ID NO. 15, or a subsequence thereof, the carrier material is used in a Southern blot. For

For long probes of at least 100 nucleotides in length, the carrier material is finally washed three times each for 15 minutes using 2 x SSC, 0.2% SDS preferably at least at 45°C (very low stringency), more preferably at least at 50°C (low stringency), more preferably at least at 55°C (medium stringency), more preferably at least at 60°C (medium-high stringency), even more preferably at least at 65°C (high stringency), and most preferably at least at 70°C (very high stringency).

For short probes which are about 15 nucleotides to about 70 nucleotides in length, stringency conditions are defined as prehybridization, hybridization, and washing post-hybridization at 5°C to 10°C below the calculated T_m using the calculation according to Bolton and McCarthy (1962, *Proceedings of the National Academy of Sciences USA* 48:1390) in 0.9 M NaCl, 0.09 M Tris-HCl pH 7.6, 6 mM EDTA, 0.5% NP-40, 1X Denhardt's solution, 1 mM sodium pyrophosphate, 1 mM sodium monobasic phosphate, 0.1 mM ATP, and 0.2 mg of yeast RNA per ml following standard Southern blotting procedures.

For short probes which are about 15 nucleotides to about 70 nucleotides in length, the carrier material is washed once in 6X SCC plus 0.1% SDS for 15 minutes and twice each for 15 minutes using 6X SSC at 5°C to 10°C below the calculated T_m.

In a third embodiment, the present invention relates to variants of the polypeptide having an amino acid sequence of SEQ ID NO. 2 or SEQ ID NO. 16, comprising a substitution, deletion, and/or insertion of one or more amino acids.

The amino acid sequences of the variant polypeptides may differ from the amino acid sequence of SEQ ID NO. 2 or SEQ ID NO. 16, or the mature polypeptide thereof, by an insertion or deletion of one or more amino acid residues and/or the substitution of one or more amino acid residues by different amino acid residues. Preferably, amino acid changes are of a minor nature, that is conservative amino acid substitutions that do not significantly affect the folding and/or activity of the protein; small deletions, typically of one to about 30 amino acids; small amino- or carboxyl-terminal extensions, such as an amino-terminal methionine residue; a small linker peptide of up to about 20-25 residues; or a small extension that facilitates purification by changing net charge or another function, such as a poly-histidine tract, an antigenic epitope or a binding domain.

Examples of conservative substitutions are within the group of basic amino acids (arginine, lysine and histidine), acidic amino acids (glutamic acid and aspartic acid),

N.H. Axelsen, J. Krøll, and B. Weeks, editors, *A Manual of Quantitative Immunoelectrophoresis*, Blackwell Scientific Publications, 1973, Chapter 11.

The antibody may also be a monoclonal antibody. Monoclonal antibodies may be prepared and used, *e.g.*, according to the methods of E. Harlow and D. Lane, editors, 1988, *Antibodies, A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring Harbor, New York.

The polypeptides of the present invention have at least 20%, preferably at least 40%, more preferably at least 60%, even more preferably at least 80%, even more preferably at least 90%, and most preferably at least 100% of the lysophospholipase activity of the mature polypeptide of SEQ ID NO. 2 or SEQ ID NO. 16.

A polypeptide of the present invention may be obtained from microorganisms of any genus. For purposes of the present invention, the term "obtained from" as used herein in connection with a given source shall mean that the polypeptide encoded by the nucleic acid sequence is produced by the source or by a cell in which the nucleic acid sequence from the source has been inserted. In a preferred embodiment, the polypeptide is secreted extracellularly.

A polypeptide of the present invention may be a bacterial polypeptide. For example, the polypeptide may be a gram positive bacterial polypeptide such as a *Bacillus* polypeptide, *e.g.*, a *Bacillus alkalophilus, Bacillus amyloliquefaciens, Bacillus brevis, Bacillus circulans, Bacillus coagulans, Bacillus lautus, Bacillus lentus, Bacillus licheniformis, Bacillus megaterium, Bacillus stearothermophilus, Bacillus subtilis*, or *Bacillus thuringiensis* polypeptide; or a *Streptomyces* polypeptide, *e.g.*, a *Streptomyces lividans* or *Streptomyces murinus* polypeptide; or a gram negative bacterial polypeptide, *e.g.*, an *E. coli* or a *Pseudomonas* sp. polypeptide.

A polypeptide of the present invention may be a fungal polypeptide, and more preferably a yeast polypeptide such as a *Candida, Kluyveromyces, Pichia, Saccharomyces, Schizosaccharomyces*, or *Yarrowia* polypeptide; or more preferably a filamentous fungal polypeptide such as an *Acremonium, Aspergillus, Aureobasidium, Cryptococcus, Filibasidium, Fusarium, Humicola, Magnaporthe, Mucor, Myceliophthora, Neocallimastix, Neurospora, Paecilomyces, Penicillium, Piromyces, Schizophyllum, Talaromyces, Thermoascus, Thielavia, Tolypocladium*, or *Trichoderma* polypeptide.

taxonomic equivalents of *Fusarium* are defined by D.L. Hawksworth, P.M. Kirk, B.C. Sutton, and D.N. Pegler (editors), 1995, *In* Ainsworth & Bisby's *Dictionary of the Fungi*, Eighth Edition, CAB International, University Press, Cambridge, England, pp. 173-174.

Strains of these species are readily accessible to the public in a number of culture collections, such as the American Type Culture Collection (ATCC), Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSM), Centraalbureau Voor Schimmelcultures (CBS), and Agricultural Research Service Patent Culture Collection, Northern Regional Research Center (NRRL).

Furthermore, such polypeptides may be identified and obtained from other sources including microorganisms isolated from nature (*e.g.*, soil, composts, water, etc.) using the above-mentioned probes. Techniques for isolating microorganisms from natural habitats are well known in the art. The nucleic acid sequence may then be derived by similarly screening a genomic or cDNA library of another microorganism. Once a nucleic acid sequence encoding a polypeptide has been detected with the probe(s), the sequence may be isolated or cloned by utilizing techniques which are known to those of ordinary skill in the art (see, *e.g.*, Sambrook *et al.*, 1989, *supra*).

As defined herein, an "isolated" polypeptide is a polypeptide which is essentially free of other non-lysophospholipase polypeptides, *e.g.*, at least about 20% pure, preferably at least about 40% pure, more preferably about 60% pure, even more preferably about 80% pure, most preferably about 90% pure, and even most preferably about 95% pure, as determined by SDS-PAGE.

Polypeptides encoded by nucleic acid sequences of the present invention also include fused polypeptides or cleavable fusion polypeptides in which another polypeptide is fused at the N-terminus or the C-terminus of the polypeptide or fragment thereof. A fused polypeptide is produced by fusing a nucleic acid sequence (or a portion thereof) encoding another polypeptide to a nucleic acid sequence (or a portion thereof) of the present invention. Techniques for producing fusion polypeptides are known in the art, and include ligating the coding sequences encoding the polypeptides so that they are in frame and that expression of the fused polypeptide is under control of the same promoter(s) and terminator.

Nucleic Acid Sequences

The present invention also relates to mutant nucleic acid sequences comprising at least one mutation in the mature polypeptide coding sequence of SEQ ID NO. 15, in which the mutant nucleic acid sequence encodes a polypeptide which consists of amino acids 17 to 648 of SEQ ID NO. 16.

The techniques used to isolate or clone a nucleic acid sequence encoding a polypeptide are known in the art and include isolation from genomic DNA, preparation from cDNA, or a combination thereof. The cloning of the nucleic acid sequences of the present invention from such genomic DNA can be effected, *e.g.*, by using the well known polymerase chain reaction (PCR) or antibody screening of expression libraries to detect cloned DNA fragments with shared structural features. See, *e.g.*, Innis *et al.*, 1990, *PCR: A Guide to Methods and Application*, Academic Press, New York. Other nucleic acid amplification procedures such as ligase chain reaction (LCR), ligated activated transcription (LAT) and nucleic acid sequence-based amplification (NASBA) may be used. The nucleic acid sequence may be cloned from a strain of *Fusarium*, or another or related organism and thus, for example, may be an allelic or species variant of the polypeptide encoding region of the nucleic acid sequence.

The term "isolated nucleic acid sequence" as used herein refers to a nucleic acid sequence which is essentially free of other nucleic acid sequences, *e.g.*, at least about 20% pure, preferably at least about 40% pure, more preferably at least about 60% pure, even more preferably at least about 80% pure, and most preferably at least about 90% pure as determined by agarose electrophoresis. For example, an isolated nucleic acid sequence can be obtained by standard cloning procedures used in genetic engineering to relocate the nucleic acid sequence from its natural location to a different site where it will be reproduced. The cloning procedures may involve excision and isolation of a desired nucleic acid fragment comprising the nucleic acid sequence encoding the polypeptide, insertion of the fragment into a vector molecule, and incorporation of the recombinant vector into a host cell where multiple copies or clones of the nucleic acid sequence will be replicated. The nucleic acid sequence may be of genomic, cDNA, RNA, semisynthetic, synthetic origin, or any combinations thereof.

The present invention also relates to nucleic acid sequences which have a degree of homology to the mature polypeptide coding sequence of SEQ ID NO. 1 (*i.e.*, nucleotides 214 to 2061) or SEQ ID NO. 15 (*i.e.*, nucleotides 49 to 1944) of at least

activity of the molecule. Sites of substrate-enzyme interaction can also be determined by analysis of the three-dimensional structure as determined by such techniques as nuclear magnetic resonance analysis, crystallography or photoaffinity labelling (see, *e.g.*, de Vos *et al.*, 1992, *Science* 255: 306-312; Smith *et al.*, 1992, *Journal of Molecular Biology* 224: 899-904; Wlodaver *et al.*, 1992, *FEBS Letters* 309: 59-64).

The present invention also relates to isolated nucleic acid sequences encoding a polypeptide of the present invention, which hybridize under very low stringency conditions, preferably low stringency conditions, more preferably medium stringency conditions, more preferably medium-high stringency conditions, even more preferably high stringency conditions, and most preferably very high stringency conditions with a nucleic acid probe which hybridizes under the same conditions with the nucleic acid sequence of SEQ ID NO. 1 or SEQ ID NO. 16, or its complementary strand; or allelic variants and subsequences thereof (Sambrook *et al.*, 1989, *supra*), as defined herein.

The present invention also relates to isolated nucleic acid sequences produced by (a) hybridizing a DNA under very low, low, medium, medium-high, high, or very high stringency conditions with (i) nucleotides 214 to 2061 of SEQ ID NO. 1 or nucleotides 49 to 1944 of SEQ ID NO. 15, (ii) the genomic DNA sequence containing nucleotides 214 to 2061 of SEQ ID NO. 1 or nucleotides 49 to 1944 of SEQ ID NO. 15, (iii) a subsequence of (i) or (ii), or (iv) a complementary strand of (i), (ii), or (iii); and (b) isolating the nucleic acid sequence. The subsequence is preferably a sequence of at least 100 nucleotides such as a sequence which encodes a polypeptide fragment which has lysophospholipase activity.

Methods for Producing Mutant Nucleic Acid Sequences

The present invention further relates to methods for producing a mutant nucleic acid sequence, comprising introducing at least one mutation into the mature polypeptide coding sequence of SEQ ID NO. 1 or SEQ ID NO. 15, or a subsequence thereof, wherein the mutant nucleic acid sequence encodes a polypeptide which consists of amino acids 38 to 654 of SEQ ID NO. 2 or amino acids 17 to 648 of SEQ ID NO. 16, or a fragment thereof which has lysophospholipase activity.

The introduction of a mutation into the nucleic acid sequence to exchange one nucleotide for another nucleotide may be accomplished by site-directed mutagenesis

An isolated nucleic acid sequence encoding a polypeptide of the present invention may be manipulated in a variety of ways to provide for expression of the polypeptide. Manipulation of the nucleic acid sequence prior to its insertion into a vector may be desirable or necessary depending on the expression vector. The techniques for modifying nucleic acid sequences utilizing recombinant DNA methods are well known in the art.

The term "control sequences" is defined herein to include all components which are necessary or advantageous for the expression of a polypeptide of the present invention. Each control sequence may be native or foreign to the nucleic acid sequence encoding the polypeptide. Such control sequences include, but are not limited to, a leader, polyadenylation sequence, propeptide sequence, promoter, signal peptide sequence, and transcription terminator. At a minimum, the control sequences include a promoter, and transcriptional and translational stop signals. The control sequences may be provided with linkers for the purpose of introducing specific restriction sites facilitating ligation of the control sequences with the coding region of the nucleic acid sequence encoding a polypeptide. The term "operably linked" is defined herein as a configuration in which a control sequence is appropriately placed at a position relative to the coding sequence of the DNA sequence such that the control sequence directs the expression of a polypeptide.

The control sequence may be an appropriate promoter sequence, a nucleic acid sequence which is recognized by a host cell for expression of the nucleic acid sequence. The promoter sequence contains transcriptional control sequences which mediate the expression of the polypeptide. The promoter may be any nucleic acid sequence which shows transcriptional activity in the host cell of choice including mutant, truncated, and hybrid promoters, and may be obtained from genes encoding extracellular or intracellular polypeptides either homologous or heterologous to the host cell.

Examples of suitable promoters for directing the transcription of the nucleic acid constructs of the present invention, especially in a bacterial host cell, are the promoters obtained from the *E. coli lac* operon, *Streptomyces coelicolor* agarase gene (*dagA*), *Bacillus subtilis* levansucrase gene (*sacB*), *Bacillus licheniformis* alpha-amylase gene (*amyL*), *Bacillus stearothermophilus* maltogenic amylase gene (*amyM*), *Bacillus amyloliquefaciens* alpha-amylase gene (*amyQ*), *Bacillus licheniformis* penicillinase gene (*penP*), *Bacillus subtilis xylA* and *xylB* genes, and prokaryotic beta-lactamase gene

and *Saccharomyces cerevisiae* glyceraldehyde-3-phosphate dehydrogenase. Other useful terminators for yeast host cells are described by Romanos *et al.*, 1992, *supra*.

The control sequence may also be a suitable leader sequence, a nontranslated region of an mRNA which is important for translation by the host cell. The leader sequence is operably linked to the 5' terminus of the nucleic acid sequence encoding the polypeptide. Any leader sequence that is functional in the host cell of choice may be used in the present invention.

Preferred leaders for filamentous fungal host cells are obtained from the genes for *Aspergillus oryzae* TAKA amylase and *Aspergillus nidulans* triose phosphate isomerase.

Suitable leaders for yeast host cells are obtained from the genes for *Saccharomyces cerevisiae* enolase (ENO-1), *Saccharomyces cerevisiae* 3-phosphoglycerate kinase, *Saccharomyces cerevisiae* alpha-factor, and *Saccharomyces cerevisiae* alcohol dehydrogenase/glyceraldehyde-3-phosphate dehydrogenase (ADH2/GAP).

The control sequence may also be a polyadenylation sequence, a sequence operably linked to the 3' terminus of the nucleic acid sequence and which, when transcribed, is recognized by the host cell as a signal to add polyadenosine residues to transcribed mRNA. Any polyadenylation sequence which is functional in the host cell of choice may be used in the present invention.

Preferred polyadenylation sequences for filamentous fungal host cells are obtained from the genes for *Aspergillus oryzae* TAKA amylase, *Aspergillus niger* glucoamylase, *Aspergillus nidulans* anthranilate synthase, *Fusarium oxysporum* trypsin-like protease, and *Aspergillus niger* alpha-glucosidase.

Useful polyadenylation sequences for yeast host cells are described by Guo and Sherman, 1995, *Molecular Cellular Biology* 15: 5983-5990.

The control sequence may also be a signal peptide coding region that codes for an amino acid sequence linked to the amino terminus of a polypeptide and directs the encoded polypeptide into the cell's secretory pathway. The 5' end of the coding sequence of the nucleic acid sequence may inherently contain a signal peptide coding region naturally linked in translation reading frame with the segment of the coding region which encodes the secreted polypeptide. Alternatively, the 5' end of the coding sequence may contain a signal peptide coding region which is foreign to the coding sequence. The

In a preferred embodiment, the propeptide coding region is nucleotides 151 to 213 of SEQ ID NO. 1 which encodes amino acids 18 to 37 of SEQ ID NO. 2.

Where both signal peptide and propeptide regions are present at the amino terminus of a polypeptide, the propeptide region is positioned next to the amino terminus of a polypeptide and the signal peptide region is positioned next to the amino terminus of the propeptide region.

It may also be desirable to add regulatory sequences which allow the regulation of the expression of the polypeptide relative to the growth of the host cell. Examples of regulatory systems are those which cause the expression of the gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound. Regulatory systems in prokaryotic systems include the *lac*, *tac*, and *trp* operator systems. In yeast, the ADH2 system or GAL1 system may be used. In filamentous fungi, the TAKA alpha-amylase promoter, *Aspergillus niger* glucoamylase promoter, and *Aspergillus oryzae* glucoamylase promoter may be used as regulatory sequences. Other examples of regulatory sequences are those which allow for gene amplification. In eukaryotic systems, these include the dihydrofolate reductase gene which is amplified in the presence of methotrexate, and the metallothionein genes which are amplified with heavy metals. In these cases, the nucleic acid sequence encoding the polypeptide would be operably linked with the regulatory sequence.

The present invention also relates to nucleic acid constructs for altering the expression of an endogenous gene encoding a polypeptide of the present invention. The constructs may contain the minimal number of components necessary for altering expression of the endogenous gene. In one embodiment, the nucleic acid constructs preferably contain (a) a targeting sequence, (b) a regulatory sequence, (c) an exon, and (d) a splice-donor site. Upon introduction of the nucleic acid construct into a cell, the construct inserts by homologous recombination into the cellular genome at the endogenous gene site. The targeting sequence directs the integration of elements (a)-(d) into the endogenous gene such that elements (b)-(d) are operably linked to the endogenous gene. In another embodiment, the nucleic acid constructs contain (a) a targeting sequence, (b) a regulatory sequence, (c) an exon, (d) a splice-donor site, (e) an intron, and (f) a splice-acceptor site, wherein the targeting sequence directs the integration of elements (a)-(f) such that elements (b)-(f) are operably linked to the

The splice-donor site of the constructs directs the splicing of one exon to another exon. Typically, the first exon lies 5' of the second exon, and the splice-donor site overlapping and flanking the first exon on its 3' side recognizes a splice-acceptor site flanking the second exon on the 5' side of the second exon. A splice-acceptor site, like a splice-donor site, is a sequence which directs the splicing of one exon to another exon. Acting in conjunction with a splice-donor site, the splicing apparatus uses a splice-acceptor site to effect the removal of an intron.

Expression Vectors

The present invention also relates to recombinant expression vectors comprising a nucleic acid sequence of the present invention, a promoter, and transcriptional and translational stop signals. The various nucleic acid and control sequences described above may be joined together to produce a recombinant expression vector which may include one or more convenient restriction sites to allow for insertion or substitution of the nucleic acid sequence encoding the polypeptide at such sites. Alternatively, the nucleic acid sequence of the present invention may be expressed by inserting the nucleic acid sequence or a nucleic acid construct comprising the sequence into an appropriate vector for expression. In creating the expression vector, the coding sequence is located in the vector so that the coding sequence is operably linked with the appropriate control sequences for expression.

The recombinant expression vector may be any vector (*e.g.*, a plasmid or virus) which can be conveniently subjected to recombinant DNA procedures and can bring about the expression of the nucleic acid sequence. The choice of the vector will typically depend on the compatibility of the vector with the host cell into which the vector is to be introduced. The vectors may be linear or closed circular plasmids.

The vector may be an autonomously replicating vector, *i.e.*, a vector which exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, *e.g.*, a plasmid, an extrachromosomal element, a minichromosome, or an artificial chromosome. The vector may contain any means for assuring self-replication. Alternatively, the vector may be one which, when introduced into the host cell, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. Furthermore, a single vector or plasmid or two or more vectors or

integrational elements may be non-encoding or encoding nucleic acid sequences. On the other hand, the vector may be integrated into the genome of the host cell by non-homologous recombination.

For autonomous replication, the vector may further comprise an origin of replication enabling the vector to replicate autonomously in the host cell in question. Examples of bacterial origins of replication are the origins of replication of plasmids pBR322, pUC19, pACYC177, and pACYC184 permitting replication in *E. coli*, and pUB110, pE194, pTA1060, and pAMß1 permitting replication in *Bacillus*. Examples of origins of replication for use in a yeast host cell are the 2 micron origin of replication, ARS1, ARS4, the combination of ARS1 and CEN3, and the combination of ARS4 and CEN6. The origin of replication may be one having a mutation which makes its functioning temperature-sensitive in the host cell (see, *e.g.*, Ehrlich, 1978, *Proceedings of the National Academy of Sciences USA* 75: 1433).

More than one copy of a nucleic acid sequence of the present invention may be inserted into the host cell to increase production of the gene product. An increase in the copy number of the nucleic acid sequence can be obtained by integrating at least one additional copy of the sequence into the host cell genome or by including an amplifiable selectable marker gene with the nucleic acid sequence where cells containing amplified copies of the selectable marker gene, and thereby additional copies of the nucleic acid sequence, can be selected for by cultivating the cells in the presence of the appropriate selectable agent.

The procedures used to ligate the elements described above to construct the recombinant expression vectors of the present invention are well known to one skilled in the art (see, *e.g.*, Sambrook *et al.*, 1989, *supra*).

Host Cells

The present invention also relates to recombinant host cells, comprising a nucleic acid sequence of the invention, which are advantageously used in the recombinant production of the polypeptides. A vector comprising a nucleic acid sequence of the present invention is introduced into a host cell so that the vector is maintained as a chromosomal integrant or as a self-replicating extra-chromosomal vector as described earlier. The term "host cell" encompasses any progeny of a parent cell that is not

and yeast belonging to the Fungi Imperfecti (Blastomycetes). Since the classification of yeast may change in the future, for the purposes of this invention, yeast shall be defined as described in *Biology and Activities of Yeast* (Skinner, F.A., Passmore, S.M., and Davenport, R.R., eds, *Soc. App. Bacteriol. Symposium Series* No. 9, 1980).

In an even more preferred embodiment, the yeast host cell is a *Candida, Hansenula, Kluyveromyces, Pichia, Saccharomyces, Schizosaccharomyces,* or *Yarrowia* cell.

In a most preferred embodiment, the yeast host cell is a *Saccharomyces carlsbergensis, Saccharomyces cerevisiae, Saccharomyces diastaticus, Saccharomyces douglasii, Saccharomyces kluyveri, Saccharomyces norbensis* or *Saccharomyces oviformis* cell. In another most preferred embodiment, the yeast host cell is a *Kluyveromyces lactis* cell. In another most preferred embodiment, the yeast host cell is a *Yarrowia lipolytica* cell.

In another more preferred embodiment, the fungal host cell is a filamentous fungal cell. "Filamentous fungi" include all filamentous forms of the subdivision Eumycota and Oomycota (as defined by Hawksworth *et al.*, 1995, *supra*). The filamentous fungi are generally characterized by a mycelial wall composed of chitin, cellulose, glucan, chitosan, mannan, and other complex polysaccharides. Vegetative growth is by hyphal elongation and carbon catabolism is obligately aerobic. In contrast, vegetative growth by yeasts such as *Saccharomyces cerevisiae* is by budding of a unicellular thallus and carbon catabolism may be fermentative.

In an even more preferred embodiment, the filamentous fungal host cell is a cell of a species of, but not limited to, *Acremonium, Aspergillus, Fusarium, Humicola, Mucor, Myceliophthora, Neurospora, Penicillium, Thielavia, Tolypocladium,* or *Trichoderma.*

In a most preferred embodiment, the filamentous fungal host cell is an *Aspergillus awamori, Aspergillus foetidus, Aspergillus japonicus, Aspergillus nidulans, Aspergillus niger* or *Aspergillus oryzae* cell. In another most preferred embodiment, the filamentous fungal host cell is a *Fusarium bactridioides, Fusarium cerealis, Fusarium crookwellense, Fusarium culmorum, Fusarium graminearum, Fusarium graminum, Fusarium heterosporum, Fusarium negundi, Fusarium oxysporum, Fusarium reticulatum, Fusarium roseum, Fusarium sambucinum, Fusarium sarcochroum, Fusarium*

sequence having at least one mutation in the mature polypeptide coding region of SEQ ID NO. 1 or SEQ ID NO. 15, wherein the mutant nucleic acid sequence encodes a polypeptide which consists of amino acids 38 to 654 of SEQ ID NO. 2 or amino acids 17 to 648 of SEQ ID NO. 16, respectively, and (b) recovering the polypeptide.

The present invention further relates to methods for producing a polypeptide of the present invention comprising (a) cultivating a homologously recombinant cell, having incorporated therein a new transcription unit comprising a regulatory sequence, an exon, and/or a splice donor site operably linked to a second exon of an endogenous nucleic acid sequence encoding the polypeptide, under conditions conducive for production of the polypeptide; and (b) recovering the polypeptide. The methods are based on the use of gene activation technology, for example, as described in U.S. Patent No. 5,641,670.

In the production methods of the present invention, the cells are cultivated in a nutrient medium suitable for production of the polypeptide using methods known in the art. For example, the cell may be cultivated by shake flask cultivation, and small-scale or large-scale fermentation (including continuous, batch, fed-batch, or solid state fermentations) in laboratory or industrial fermentors performed in a suitable medium and under conditions allowing the polypeptide to be expressed and/or isolated. The cultivation takes place in a suitable nutrient medium comprising carbon and nitrogen sources and inorganic salts, using procedures known in the art. Suitable media are available from commercial suppliers or may be prepared according to published compositions (*e.g.*, in catalogues of the American Type Culture Collection). If the polypeptide is secreted into the nutrient medium, the polypeptide can be recovered directly from the medium. If the polypeptide is not secreted, it can be recovered from cell lysates.

The polypeptides may be detected using methods known in the art that are specific for the polypeptides. These detection methods may include use of specific antibodies, formation of an enzyme product, or disappearance of an enzyme substrate. For example, an enzyme assay may be used to determine the activity of the polypeptide as described herein.

The resulting polypeptide may be recovered by methods known in the art. For example, the polypeptide may be recovered from the nutrient medium by conventional procedures including, but not limited to, centrifugation, filtration, extraction, spray-

plant or plant cell is constructed by incorporating one or more expression constructs encoding a polypeptide of the present invention into the plant host genome and propagating the resulting modified plant or plant cell into a transgenic plant or plant cell.

Conveniently, the expression construct is a nucleic acid construct which comprises a nucleic acid sequence encoding a polypeptide of the present invention operably linked with appropriate regulatory sequences required for expression of the nucleic acid sequence in the plant or plant part of choice. Furthermore, the expression construct may comprise a selectable marker useful for identifying host cells into which the expression construct has been integrated and DNA sequences necessary for introduction of the construct into the plant in question (the latter depends on the DNA introduction method to be used).

The choice of regulatory sequences, such as promoter and terminator sequences and optionally signal or transit sequences is determined, for example, on the basis of when, where, and how the polypeptide is desired to be expressed. For instance, the expression of the gene encoding a polypeptide of the present invention may be constitutive or inducible, or may be developmental, stage or tissue specific, and the gene product may be targeted to a specific tissue or plant part such as seeds or leaves. Regulatory sequences are, for example, described by Tague *et al.*, 1988, *Plant Physiology* 86: 506.

For constitutive expression, the 35S-CaMV promoter may be used (Franck *et al.*, 1980, *Cell* 21: 285-294). Organ-specific promoters may be, for example, a promoter from storage sink tissues such as seeds, potato tubers, and fruits (Edwards & Coruzzi, 1990, *Ann. Rev. Genet.* 24: 275-303), or from metabolic sink tissues such as meristems (Ito *et al.*, 1994, *Plant Mol. Biol.* 24: 863-878), a seed specific promoter such as the glutelin, prolamin, globulin, or albumin promoter from rice (Wu *et al.*, 1998, *Plant and Cell Physiology* 39: 885-889), a *Vicia faba* promoter from the legumin B4 and the unknown seed protein gene from *Vicia faba* (Conrad *et al.*, 1998, *Journal of Plant Physiology* 152: 708-711), a promoter from a seed oil body protein (Chen *et al.*, 1998, Plant and Cell Physiology 39: 935-941), the storage protein napA promoter from *Brassica napus*, or any other seed specific promoter known in the art, *e.g.*, as described in WO 91/14772. Furthermore, the promoter may be a leaf specific promoter such as the *rbcs* promoter from rice or tomato (Kyozuka *et al.*, 1993, *Plant Physiology* 102: 991-

The present invention also relates to methods for producing a polypeptide of the present invention comprising (a) cultivating a transgenic plant or a plant cell comprising a nucleic acid sequence encoding a polypeptide having lysophospholipase activity of the present invention under conditions conducive for production of the polypeptide; and (b) recovering the polypeptide.

Removal or Reduction of Lysophospholipase Activity

The present invention also relates to methods for producing a mutant cell of a parent cell, which comprises disrupting or deleting a nucleic acid sequence encoding the polypeptide or a control sequence thereof, which results in the mutant cell producing less of the polypeptide than the parent cell when cultivated under the same conditions.

The construction of strains which have reduced lysophospholipase activity may be conveniently accomplished by modification or inactivation of a nucleic acid sequence necessary for expression of the polypeptide having lysophospholipase activity in the cell. The nucleic acid sequence to be modified or inactivated may be, for example, a nucleic acid sequence encoding the polypeptide or a part thereof essential for exhibiting lysophospholipase activity, or the nucleic acid sequence may have a regulatory function required for the expression of the polypeptide from the coding sequence of the nucleic acid sequence. An example of such a regulatory or control sequence may be a promoter sequence or a functional part thereof, *i.e.*, a part which is sufficient for affecting expression of the polypeptide. Other control sequences for possible modification are described above.

Modification or inactivation of the nucleic acid sequence may be performed by subjecting the cell to mutagenesis and selecting or screening for cells in which the lysophospholipase producing capability has been reduced. The mutagenesis, which may be specific or random, may be performed, for example, by use of a suitable physical or chemical mutagenizing agent, by use of a suitable oligonucleotide, or by subjecting the DNA sequence to PCR generated mutagenesis. Furthermore, the mutagenesis may be performed by use of any combination of these mutagenizing agents.

Examples of a physical or chemical mutagenizing agent suitable for the present purpose include ultraviolet (UV) irradiation, hydroxylamine, N-methyl-N'-nitro-N-nitrosoguanidine (MNNG), O-methyl hydroxylamine, nitrous acid, ethyl methane

translated is thus reduced or eliminated.

It is preferred that the cell to be modified in accordance with the methods of the present invention is of microbial origin, for example, a fungal strain which is suitable for the production of desired protein products, either homologous or heterologous to the cell.

The present invention further relates to a mutant cell of a parent cell which comprises a disruption or deletion of a nucleic acid sequence encoding the polypeptide or a control sequence thereof, which results in the mutant cell producing less of the polypeptide than the parent cell.

The polypeptide-deficient mutant cells so created are particularly useful as host cells for the expression of homologous and/or heterologous polypeptides. Therefore, the present invention further relates to methods for producing a homologous or heterologous polypeptide comprising (a) cultivating the mutant cell under conditions conducive for production of the polypeptide; and (b) recovering the polypeptide. The term "heterologous polypeptides" is defined herein as polypeptides which are not native to the host cell, a native protein in which modifications have been made to alter the native sequence, or a native protein whose expression is quantitatively altered as a result of a manipulation of the host cell by recombinant DNA techniques.

In a further aspect, the present invention relates to a method for producing a protein product essentially free of lysophospholipase activity by fermentation of a cell which produces both a polypeptide of the present invention as well as the protein product of interest by adding an effective amount of an agent capable of inhibiting lysophospholipase activity to the fermentation broth before, during, or after the fermentation has been completed, recovering the product of interest from the fermentation broth, and optionally subjecting the recovered product to further purification.

In a further aspect, the present invention relates to a method for producing a protein product essentially free of lysophospholipase activity by cultivating the cell under conditions permitting the expression of the product, subjecting the resultant culture broth to a combined pH and temperature treatment so as to reduce the lysophospholipase activity substantially, and recovering the product from the culture broth. Alternatively, the combined pH and temperature treatment may be performed on an enzyme preparation recovered from the culture broth. The combined pH and temperature treatment may

Uses

The present invention is also directed to methods for using the polypeptides having lysophospholipase activity.

The polypeptides of the present invention may be used in any application where it is desired to hydrolyze the fatty acyl group(s) of a phospholipid or lysophospholipid, such as lecithin or lysolecithin. The polypeptides of the present invention are preferably used at a pH optimal for activity.

A polypeptide having lysophospholipase activity of the present invention may be used for degumming an aqueous carbohydrate solution or slurry to improve its filterability, particularly, a starch hydrolysate, especially a wheat starch hydrolysate which is difficult to filter and yields cloudy filtrates. The treatment may be performed using methods well known in the art. See, for example, EP 219,269 and EP 808,903.

A polypeptide having lysophospholipase activity of the present invention may be used in a process to reduce the phospholipid content in an edible oil by treating the oil with the polypeptide to hydrolyze a major portion of the phospholipid and separating an aqueous phase containing the hydrolyzed phospholipid from the oil. Such a process is applicable to the purification of any edible oil which contains phospholipid, *e.g.*, vegetable oil such as soybean oil, rape seed oil, and sunflower oil.

Prior to the lysophospholipase treatment, the oil is preferably pretreated to remove slime (mucilage), *e.g.*, by wet refining. Typically, the oil will contain 50-250 ppm of phosphorus as phospholipid at the beginning of the treatment with the lysophospholipase, and the treatment may reduce the phosphorus value to below 5-10 ppm.

The lysophospholipase treatment is conducted by dispersing an aqueous solution of the lysophospholipase, preferably as droplets with an average diameter below 10 μm. The amount of water is preferably 0.5-5% by weight in relation to the oil. An emulsifier may optionally be added. Mechanical agitation may be applied to maintain the emulsion.

The lysophospholipase treatment can be conducted at a pH in the range of about 1.5 to about 5.0. The process pH may be in the range of about 3.5 to about 5 to maximize the enzyme's performance, or a pH in the range of about 1.5 to about 3 (*e.g.*, 2-3) may be used in order to suppress the alkaline hydrolysis of triglycerides (saponification). The pH may be adjusted by adding citric acid, a citrate buffer, or

at least one property of interest of the dough and/or baked product.

The term "improved property" is defined herein as any property of a dough and/or a product obtained from the dough, particularly a baked product, which is improved by the action of the lysophospholipase activity relative to a dough or product in which the polypeptide having lysophospholipase activity is not incorporated. The improved property may include, but is not limited to, increased strength of the dough, increased elasticity of the dough, increased stability of the dough, reduced stickiness of the dough, improved extensibility of the dough, improved machinability of the dough, increased volume of the baked product, improved crumb structure of the baked product, improved softness of the baked product, improved flavor of the baked product, and/or improved antistaling of the baked product.

The improved property may be determined by comparison of a dough and/or a baked product prepared with and without addition of a polypeptide having lysophospholipase activity of the present invention in accordance with the methods of the present invention. Techniques which can be used to determine improvements achieved by use of the methods of present invention are described below in the Examples. Organoleptic qualities may be evaluated using procedures well established in the baking industry, and may include, for example, the use of a panel of trained taste-testers.

The term “increased strength of the dough” is defined herein as the property of a dough that has generally more elastic properties and/or requires more work input to mould and shape.

The term “increased elasticity of the dough” is defined herein as the property of a dough which has a higher tendency to regain its original shape after being subjected to a certain physical strain.

The term “increased stability of the dough” is defined herein as the property of a dough that is less susceptible to mechanical abuse thus better maintaining its shape and volume.

The term “reduced stickiness of the dough” is defined herein as the property of a dough that has less tendency to adhere to surfaces, *e.g.*, in the dough production machinery, and is either evaluated empirically by the skilled test baker or measured by the use of a texture analyzer (*e.g.*, TAXT2) as known in the art.

The term “improved extensibility of the dough” is defined herein as the property

particular a non-dusting granulate, liquid, in particular a stabilized liquid, or protected enzyme. Granulates and agglomerated powders may be prepared by conventional methods, *e.g.*, by spraying the lysophospholipase onto a carrier in a fluid-bed granulator. The carrier may consist of particulate cores having a suitable particle size. The carrier may be soluble or insoluble, *e.g.*, a salt (such as NaCl or sodium sulfate), sugar (such as sucrose or lactose), sugar alcohol (such as sorbitol), starch, rice, corn grits, or soy. The lysophospholipase and/or additional enzymes may be contained in slow-release formulations. Methods for preparing slow-release formulations are well known in the art. Liquid enzyme preparations may, for instance, be stabilized by adding nutritionally acceptable stabilizers such as a sugar, sugar alcohol, or another polyol, and/or lactic acid or another organic acid according to established methods.

For inclusion in pre-mixes or flour it is advantageous that the polypeptide having lysophospholipase activity is in the form of a dry product, *e.g.*, a non-dusting granulate, whereas for inclusion together with a liquid it is advantageously in a liquid form.

One or more additional enzymes may also be incorporated into the dough. The additional enzyme may be of any origin, including mammalian and plant, and preferably of microbial (bacterial, yeast or fungal) origin and may be obtained by techniques conventionally used in the art.

In a preferred embodiment, the additional enzyme may be an amylase, such as an alpha-amylase (useful for providing sugars fermentable by yeast and retarding staling) or beta-amylase, cyclodextrin glucanotransferase, peptidase, in particular, an exopeptidase (useful in flavour enhancement), transglutaminase, lipase (useful for the modification of lipids present in the dough or dough constituents so as to soften the dough), phospholipase (useful for the modification of lipids present in the dough or dough constituents so as to soften the dough and improve gas retention in the dough), cellulase, hemicellulase, in particular a pentosanase such as xylanase (useful for the partial hydrolysis of pentosans which increases the extensibility of the dough), protease (useful for gluten weakening in particular when using hard wheat flour), protein disulfide isomerase, *e.g.*, a protein disulfide isomerase as disclosed in WO 95/00636, glycosyltransferase, peroxidase (useful for improving the dough consistency), laccase, or oxidase, *e.g.*, an aldose oxidase, glucose oxidase, pyranose oxidase, lipoxygenase, or L-amino acid oxidase (useful in improving dough consistency).

consisting of amino acids 1 to 17 of SEQ ID NO. 2 or nucleotides 1 to 48 of SEQ ID NO. 15 encoding a signal peptide consisting of amino acids 1 to 16 of SEQ ID NO. 16, and a second nucleic acid sequence consisting of nucleotides 151 to 213 of SEQ ID NO. 1 encoding a propeptide consisting of amino acids 18 to 37 of SEQ ID NO. 2, wherein the gene is foreign to the first and second nucleic acid sequences.

The present invention also relates to recombinant expression vectors and recombinant host cells comprising such nucleic acid constructs.

The present invention also relates to methods for producing a protein comprising (a) cultivating such a recombinant host cell under conditions suitable for production of the protein; and (b) recovering the protein.

The first and second nucleic acid sequences may be operably linked to foreign genes individually with other control sequences or in combination with other control sequences. Such other control sequences are described *supra*. As noted earlier, where both signal peptide and propeptide regions are present at the amino terminus of a protein, the propeptide region is positioned next to the amino terminus of a protein and the signal peptide region is positioned next to the amino terminus of the propeptide region.

The protein may be native or heterologous to a host cell. The term "protein" is not meant herein to refer to a specific length of the encoded product and, therefore, encompasses peptides, oligopeptides, and proteins. The term "protein" also encompasses two or more polypeptides combined to form the encoded product. The proteins also include hybrid polypeptides which comprise a combination of partial or complete polypeptide sequences obtained from at least two different proteins wherein one or more may be heterologous or native to the host cell. Proteins further include naturally occurring allelic and engineered variations of the above mentioned proteins and hybrid proteins.

Preferably, the protein is a hormone or variant thereof, enzyme, receptor or portion thereof, antibody or portion thereof, or reporter. In a more preferred embodiment, the protein is an oxidoreductase, transferase, hydrolase, lyase, isomerase, or ligase. In an even more preferred embodiment, the protein is an aminopeptidase, amylase, carbohydrase, carboxypeptidase, catalase, cellulase, chitinase, cutinase, cyclodextrin glycosyltransferase, deoxyribonuclease, esterase, alpha-galactosidase, beta-galactosidase, glucoamylase, alpha-glucosidase, beta-glucosidase, invertase, laccase,

YEPG medium was composed per liter of 10 g of yeast extract, 20 g of peptone, and 20 g of glucose.

STC was composed of 0.8 M sorbitol, 25 mM Tris pH 8, 25 mM $CaCl_2$.

SPTC was composed of 40% PEG 4000, 0.8 M sorbitol, 25 mM Tris pH 8, 25 mM $CaCl_2$.

M400Da medium was composed per liter of 50 g of maltodextrin, 2 g of $MgSO_4 \cdot 7H_2O$, 2 g of KH_2PO_4, 4 g of citric acid, 8 g of yeast extract, 2 g of urea, and 1 ml of COVE trace metals solution.

10X basal salts w/o amino acid was composed per liter of 66.8 g of yeast nitrogen base w/o amino acids (Diffco), 100 g of succinic acid, and 60 g of NaOH.

SC Ura-glc medium was composed per liter of 100 ml of 20% glucose, 100 ml of 10X basal salts w/o amino acid, 25 ml of 20% (w/v) casamino acid, 4 ml of 5% of threonine, 10 ml of 1% tryptophan, and 20 g of Agar Noble.

SC Ura-gal medium was composed per liter of 100 ml of 20% galactose, 100 ml of 10X basal salts w/o amino acid , 25 ml of 20% (w/v) casamino acid, 4 ml of 5% of threonine, 10 ml of 1% tryptophan and 20 g of Noble agar.

YPD medium was composed per liter of 10 g of yeast extract, 20 g of bactopeptone, and 100 ml of 20% glucose.

1X TE/LiAc was composed per 10 ml of 1ml of 10X TE (100 mM of Tris and 10 mM of EDTA at pH), 1 ml of 1 M lithium acetate, and 8 ml of milli Q water.

PEG/LiAc solution was composed of 50 ml of 40% of PEG and 1 ml of 5 M lithium acetate.

Assay plates for detecting lysophospholipase activity were composed per liter of 5 g of L-alpha phosphatidylcholine 95%, 2.5 g of cholic acid, 50 ml of 1 M Tris-HCl pH 8.0 buffer, 100 ml of 100 mM $CaCl_2$, 15 ml of 2% Brilliant green, and 20 g of Agar Noble. The solution was poured into Falcon 1058 plates in 50 ml aliquots.

STET solution was composed of 8% of sucrose, 50 mM Tris-HCl pH 8.0, 50mM EDTA, and 5% of Triton X-100.

Example 1: Fermentation and Mycelial Tissue

Fusarium venenatum CC1-3, a morphological mutant of *Fusarium* strain ATCC 20334 (Wiebe *et al.*, 1991, *Mycol. Research* 95: 1284-1288), was grown in a two-liter

cloning vector pZErO-2.1. Library A was made using mRNA from mycelia harvested at four days, and Library B was constructed with mRNA from the six day time point. Neither cDNA library was amplified in order to examine a representative "snapshot" of the gene expression profile in the cells. Instead the libraries were plated, titered, and independent clones from each was analyzed by DNA sequencing.

Library A (4 day cells) consisted about 7.5 x 10^4 independent clones and Library B (6 day cells) consisted of roughly 1.2 x 10^5 clones. Miniprep DNA was isolated from forty colonies in each library and checked for the presence and size of cDNA inserts. In this analysis 39 of 40 colonies (97.5%) from Library A contained inserts with sizes ranging from 600 bp to 2200 bp (avg. = 1050 bp). Similarly, 39 of 40 colonies (97.5%) picked from Library B had inserts with sizes ranging from 800 bp to 3600 bp (avg. = 1380 bp).

Example 3: Template Preparation and Nucleotide Sequencing

From each cDNA library described in Example 2, 1192 transformant colonies were picked directly from the transformation plates into 96-well microtiter dishes which contained 200 µl of 2YT broth (Miller, 1992, *supra)* with 50 µg/ml kanamycin. The plates were incubated overnight at 37°C without shaking. After incubation 100 µl of sterile 50% glycerol was added to each well. The transformants were replicated into secondary, deep-dish 96-well microculture plates (Advanced Genetic Technologies Corporation, Gaithersburg, MD) containing 1 ml of Magnificent Broth™ (MacConnell Research, San Diego, CA) supplemented with 50 µg of kanamycin per ml in each well. The primary microtiter plates were stored frozen at -80°C. The secondary deep-dish plates were incubated at 37°C overnight with vigorous agitation (300 rpm) on rotary shaker. To prevent spilling and cross-contamination, and to allow sufficient aeration, each secondary culture plate was covered with a polypropylene pad (Advanced Genetic Technologies Corporation, Gaithersburg, MD) and a plastic microtiter dish cover.

DNA was isolated from each well using the 96-well Miniprep Kit protocol of Advanced Genetic Technologies Corporation (Gaithersburg, MD) as modified by Utterback *et al.* (1995, *Genome Sci. Technol.* 1: 1-8). Single-pass DNA sequencing was done with a Perkin-Elmer Applied Biosystems Model 377 Sequencer XL (Perkin-Elmer Applied Biosystems, Inc., Foster City, CA) using dye-terminator chemistry (Giesecke *et*

Automated DNA Sequencer using dye-terminator chemistry. Contiguous sequences were generated using a transposon insertion strategy (Primer Island Transposition Kit, Perkin-Elmer/Applied Biosystems, Inc., Foster City, CA). The lysophospholipase clone from *E. coli* FB0346 was sequenced to an average redundancy of 6.9.

The lysophospholipase clone encoded an open reading frame of 1962 bp encoding a polypeptide of 654 amino acids. The nucleotide sequence (SEQ ID NO. 1) and deduced amino acid sequence (SEQ ID NO. 2) are shown in Figures 1A, 1B, and 1C. Using the SignalP program (Nielsen *et al.*, 1997, *Protein Engineering* 10: 1-6), a signal peptide of 17 residues was predicted and N-terminal analysis of the secreted protein indicated the presence of a pro region of 20 amino acids (see Example 11), hence indicating a predicted molecular weight of approximately 67 kDa for the secreted lysophospholipase. Thus, the mature lysophospholipase is composed of 617 amino acids.

A comparative alignment of lysophospholipase sequences was undertaken using the Clustal method (Higgins, 1989, *CABIOS* 5: 151-153) using the LASERGENE™ MEGALIGN™ software (DNASTAR, Inc., Madison, WI) with an identity table and the following multiple alignment parameters: Gap penalty of 10, and gap length penalty of 10. Pairwise alignment parameters were Ktuple=1, gap penalty=3, windows=5, and diagonals=5.

The comparative alignment showed that the *Fusarium venenatum* lysophospholipase shares regions of identity with lysophospholipase proteins from *Neurospora crassa* of 59% (TREMBL O42791), *Penicillium notatum* of 52% (Swissprot P39457), *Saccharomyces cerevisiae* of 44% (Swissprot P39105), and *Schizosaccharomyces pombe* of 39% (TREMBL O13857). The identities are highest between regions that are likely to be important for catalytic and/or structural roles of the enzyme. There are 19 potential N-linked glycosylation sites (Asn-X-Ser/Thr) within *Fusarium venenatum* lysophospholipase, and 11 of these are conserved in *Neurospora crassa* lysophospholipase, whereas 10 are conserved in *Penicillium notatum* lysophospholipase. The alignment also indicates the presence of eight Cys residues whose positions are strictly conserved among *Fusarium venenatum, Neurospora crassa, Penicillium notatum, Saccharomyces pombe*, and *Saccharomyces cerevisiae* lysophospholipases.

5)

Primer 4 (antisense):

5'-TGGTATGCATAAGCTTGAATTCAGGTAAACAAGATATAATTT-3' (SEQ ID NO. 6)

*Nsi*I *Hind*III *Eco*RI

The final 2.3 kb overlapping PCR fragment which contained -1208 to -1 of the *Fusarium oxysporum* trypsin promoter, the 25 base pair polylinker and 1060 base pairs of the *Fusarium oxysporum* trypsin terminator was obtained using 0.2 µl of the first PCR (promoter) reaction and 3 µl of the second (terminator) reaction as templated and primers 1 and 4. The PCR conditions used were 95°C for 3 minutes followed by 30 cycles each at 95°C for 30 seconds, 62°C for 1 minute, and 72°C for 3 minutes. The final extension cycle was at 72°C for 5 minutes. *Pwo* DNA polymerase was also used for this reaction.

The resulting 2.3 kb fragment containing the trypsin promoter, the polylinker, and the trypsin terminator was digested with *Eco*RI and ligated into the *Eco*RI digested vector pMT1612 containing the *bar* gene (WO 97/26330) to create pDM181 (Figure 2).

Example 8: Construction of plasmid pSheB1

The *Fusarium venenatum* expression vector pSheB1 (Figure 3) was generated by modification of pDM181. The modifications included (a) removal of two *Nco*I sites within the pDM181 sequence, and (b) restoration of the natural translation start of the Fusarium oxysporum trypsin promoter (reconstruction of an *Nco*I site at the ATG start codon).

Removal of two *Nco*I sites within the pDM181 sequence was accomplished using the QuikChange™ site-directed mutagenesis kit (Stratagene Cloning Systems, La Jolla, CA) according to the manufacturer's instruction with the following pairs of mutagenesis primers:

5'-dCAGTGAATTGGCCTCGATGGCCGCGGCCGCGAATT-3' plus (SEQ ID NO. 7)

5'-dAATTCGCGGCCGCGGCCATCGAGGCCAATTCACTG-3' (SEQ ID NO. 8)

5'-dCACGAAGGAAAGACGATGGCTTTCACGGTGTCTG-3' plus (SEQ ID NO. 9)

5'-dCAGACACCGTGAAAGCCATCGTCTTTCCTTCGTG-3' (SEQ ID NO. 10)

Restoration of the natural translation start of the *Fusarium oxysporum* trypsin promoter was also accomplished using the Stratagene QuikChange™ site directed

Spores of *Fusarium venenatum* CC1-3 (MLY-3) were generated by inoculating a flask containing 500 ml of RA sporulation medium with 10 plugs from a 1X Vogels medium plate (2.5% Noble agar) supplemented with 2.5% glucose and 2.5 mM sodium nitrate and incubating at 28°C, 150 rpm for 2 to 3 days. Spores were harvested through Miracloth (Calbiochem, San Diego, CA) and centrifuged 20 minutes at 7000 rpm in a Sorvall RC-5B centrifuge (E. I. DuPont De Nemours and Co., Wilmington, DE). Pelleted spores were washed twice with sterile distilled water, resuspended in a small volume of water, and then counted using a hemocytometer.

Protoplasts were prepared by inoculating 100 ml of YEPG medium with 4 X 10^7 spores of *Fusarium venenatum* CC1-3 and incubating for 16 hours at 24°C and 150 rpm. The culture was centrifuged for 7 minutes at 3500 rpm in a Sorvall RT 6000D (E. I. DuPont De Nemours and Co., Wilmington, DE). Pellets were washed twice with 30 ml of 1 M $MgSO_4$ and resuspended in 15 ml of 5 mg/ml of NOVOZYME 234™ (batch PPM 4356, Novo Nordisk A/S, Bagsværd, Denmark) in 1 M $MgSO_4$. Cultures were incubated at 24°C and 150 rpm until protoplasts formed. A volume of 35 ml of 2 M sorbitol was added to the protoplast digest and the mixture was centrifuged at 2500 rpm for 10 minutes. The pellet was resuspended, washed twice with STC, and centrifuged at 2000 rpm for 10 minutes to pellet the protoplasts. Protoplasts were counted with a hemocytometer and resuspended in an 8:2:0.1 solution of STC:SPTC:DMSO to a final concentration of 1.25 x 10^7 protoplasts/ml. The protoplasts were stored at -80°C, after controlled-rate freezing in a Nalgene Cryo 1°C Freezing Container (VWR Scientific, Inc., San Francisco, CA).

Frozen protoplasts of *Fusarium venenatum* CC1-3 were thawed on ice. Five µg of pRaMB54 described in Example 9 and 5 µl of heparin (5 mg per ml of STC) was added to a 50 ml sterile polypropylene tube. One hundred µl of protoplasts was added, mixed gently, and incubated on ice for 30 minutes. One ml of SPTC was added and incubated 20 minutes at room temperature. After the addition of 25 ml of 40°C COVE top agarose, the mixture was poured onto an empty 150 mm diameter plate and incubated overnight at room temperature. Then an additional 25 ml of 40°C COVE top agarose containing 10 mg of BASTA™ per ml was poured on top of the plate and incubated at room temperature for up to 14 days. The active ingredient in the herbicide BASTA™ is phosphinothricin. BASTA™ was obtained from AgrEvo (Hoechst Schering, Rodovre,

Table 1. Lysophospholipase activity present in culture supernatants from *Fusarium venenatum*/pRaMB54 transformants.

Transformant	Relative Lysophospholipase Activity†
RaMB54.01	1.00
RaMB54.02	0.86
RaMB54.03	0.88
RaMB54.04	0.88
RaMB54.05	0.96
RaMB54.06	0.08
RaMB54.10	0.85
RaMB54.11	0.88
RaMB54.12	0.88
RaMB54.14	0.35
RaMB54.15	1.00
RaMB54.16	0.88
RaMB54.17	0.88
RaMB54.18	0.81
RaMB54.19	0.85
MLY-3 control	N.D.

†Activity reflects the rate of hydrolysis of egg yolk lysolecithin at pH 7 and 37°C measured in micromoles of product per minute per ml relative to the activity of the highest producer RamB54.01 which is normalized to 1.00. N.D., not detected.

Example 11: Purification of recombinant *Fusarium venenatum* lysophospholipase

Fusarium venenatum/pRaMB54 was cultivated as described in Example 10 for 4 days in two 500 ml shake flasks containing 100 ml of M400Da medium. The 4 day whole culture broths were filtered through Miracloth followed by a 0.45 μm syringe filter (Whatman, Inc., Fairfield, NJ) to yield a sample volume of approximately 150 ml. Then 25 mM PMSF in 75% ethanol/25% methanol was added to the filtered broth to a final concentration of 0.5 mM. The sample was then diluted with water and 20 mM sodium phosphate pH 7 to achieve a pH of 7.45 and a conductivity of 2.3 mS.

volts in 10 mM CAPS pH 11.0 buffer. The PVDF membranes were stained in 0.1% Commassie Blue R250 in 40% methanol/1% acetic acid and the observed bands excised. The excised bands were sequenced from a blot cartridge using sequencing reagents (Perkin Elmer/Applied Biosystems Division, Foster City, CA). Detection of phenylthiohydantoin-amino acids was accomplished by on-line HPLC using Buffer A containing 3.5% tetrahydrofuran in water with 18 ml of the Premix concentrate (Perkin Elmer/Applied Biosystems Division, Foster City, CA) containing acetic acid, sodium acetate, and sodium hexanesulfonate and Buffer B containing acetonitrile. Data was collected and analyzed on a Macintosh IIsi using Applied Biosystems 610 Data Analysis software.

SDS-PAGE of the purified lysophospholipase revealed major bands at approximately 130 kDa and 90 kDa, and minor bands at approximately 45, 40, and 35 kDa. Molecular weights were based on pre-stained Multi-Mark SDS page markers which do not reflect accurate MW determinations. N-terminal sequencing of the excised bands produced the following sequences.

Run# AB0909 - 130 kDa band sequence: ALPDSPSGGY (SEQ ID NO. 2).

Run# AB0910 - 90 kDa band sequence: NTAKYWDDIKDTVDEKADGW (SEQ ID NO. 2) (Internal peptide following leucine).

Run# AB0912 - 45 kDa band sequence: ALPDSPSGGYA (SEQ ID NO. 2)

Run# AB0913 - 40 kDa band sequence: ALPDSPSGGYAPKV (SEQ ID NO. 2)

Run# AB0915 - 35 kDa band sequence: ALPD(S)P?GGYAP (SEQ ID NO. 2).

SDS-PAGE of the unpurified lysophospholipase revealed a major band at 116 kDa based on Novex Mark 12 SDS-page markers. N-terminal sequencing of the excised 116 kDa band produced the following sequence:

Run# AB0917 - 116 kDa band sequence: ALPDSPSGGYAPKVVD?P (SEQ ID NO. 2) (?=Cys which is undetected without modification in Edman chemistry). The N-terminus appeared to be processed with a 20 amino acid pro-peptide. Cleavage occurred following an arginine.

The overall results are summarized as follows:

Signal Sequence: MLGPLVFTLWLTSSAIA (SEQ ID NO. 2)

A Grain Fatty Acid Methyl Ester Mix (Supelco, Inc., Bellefonte, PA) was used to determine the retention time for lauroyl fatty acid methyl ester and also for quantitation of the fatty acid peaks.

The results are shown in Figure 5. No hydrolysis of the substrate was detected at pH 2 and 8. The pH optimum was determined to be ~ 4-6.

Example 14: Construction of *Fusarium verticillioides* lysophospholipase cDNA library

Mycelia of *Fusarium verticillioides* were prepared similarly as described in Example 1. Poly A^+ mRNA was isolated from mycelia of *Fusarium verticillioides* and cDNA was synthesized from the Poly A^+ mRNA with *Bst*XI and *Not*I linkers similarly as described in Example 2. The cDNA cloned into the yeast expression vector pYES 2.0 (Invitrogen Corporation, Carlsbad, CA) according to the manufacturer's instructions.

Example 15: Transformation of *Saccharomyces cerevisiae* with *Fusarium verticillioides* lysophospholipase cDNA library

Competent cells of *Saccharomyces cerevisiae* YNG318 were prepared using the following procedure. *Saccharomyces cerevisiae* YNG318 was grown in 20 ml of YPD medium overnight at 30°C. A sufficient seed culture was transferred to 300 ml of YPD medium and cultivated at 30°C for 3 hours at 230-250 rpm until the OD_{600} = 0.2-0.3. Then the cells were harvested by centrifugation at 5000 rpm for 5 minutes at 20°C. The pellet was suspended in 50 ml of sterile water and the suspension was centrifuged again. The pellet was then suspended in 1.5 ml of 1X TE/LiAc and glycerol was added to a final concentration of 15%. The competent cells were stored at -80°C until use.

Transformation of *Saccharomyces cerevisiae* YNG318 was performed with the *Fusarium verticillioides* cDNA library described in Example 14 using the following procedure. A 100 µl volume of competent cells was thawed on ice and transferred into a sterile tube containing 10 µl of carrier DNA (yeast marker carrier DNA; Clontech Laboratories, Inc., Palo Alto, CA). One µg of plasmid DNA from Example 14 was added to the tube and the tube was gently mixed. Then 0.6 ml of sterile PEG/LiAc solution was added to the tube. The tube was incubated at 30°C with shaking at 200 rpm for 30 minutes and then heat-shocked at 42°C for 15 minutes. After heat-shock the tube was

the Clustal method (Higgins, 1989, *CABIOS* 5: 151-153) using the LASERGENE™ MEGALIGN™ software (DNASTAR, Inc., Madison, WI) with an identity table and the following multiple alignment parameters: Gap penalty of 10, and gap length penalty of 10. Pairwise alignment parameters were Ktuple=1, gap penalty=3, windows=5, and diagonals=5.

The comparative alignment showed that the *Fusarium verticillioides* lysophospholipase shares regions of identity with phospholipase proteins *Neurospora crassa* of 55.4% (042791), *Penicillium notatum* of 49.3% (Swissprot P39457), *Saccharomyces cerevisiae* of 37.0% (EMBL L23089), and *Torulaspora delbrueckii* of 38.3% (EMBL D32134). The identities are highest between regions that are likely to be important for catalytic and/or structural roles of the enzyme.

Deposit of Biological Materials

The following biological materials have been deposited under the terms of the Budapest Treaty with the Agricultural Research Service Patent Culture Collection, Northern Regional Research Center, 1815 University Street, Peoria, Illinois, 61604, and Centraalbureau voor Schimmelcultures, Oosterstraat 1, P.O. Box 273, 3740 AG Baarn, The Netherlands, and given the following accession numbers:

Deposit	Accession Number	Date of Deposit
E. coli TOP10 (pFB0346)	NRRL B-30073	October 27, 1998
Fusarium verticillioides	CBS 650.96	June 5, 1996

The invention described and claimed herein is not to be limited in scope by the specific embodiments herein disclosed, since these embodiments are intended as illustrations of several aspects of the invention. Any equivalent embodiments are intended to be within the scope of this invention. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description. Such modifications are also intended to fall within the scope of the appended claims. In the case of conflict, the present disclosure including definitions will control.

Various references are cited herein, the disclosures of which are incorporated by reference in their entireties.

90% identity with amino acids 38 to 654 of SEQ ID NO. 2 or amino acids 17 to 648 of SEQ ID NO. 16.

6. The polypeptide of claim 5, having an amino acid sequence which has at least 95% identity with amino acids 38 to 654 of SEQ ID NO. 2 or amino acids 17 to 648 of SEQ ID NO. 16.

7. The polypeptide of any of claims 1-6, comprising the amino acid sequence of SEQ ID NO. 2 or SEQ ID NO. 16.

8. The polypeptide of any of claims 1-7, consisting of the amino acid sequence of SEQ ID NO. 2 or SEQ ID NO. 16, or a fragment thereof.

9. The polypeptide of claim 8, consisting of the amino acid sequence of SEQ ID NO. 2 or SEQ ID NO. 16.

10. The polypeptide of claim 9, which consists of amino acids 38 to 654 of SEQ ID NO. 2 or amino acids 17 to 648 of SEQ ID NO. 16.

11. The polypeptide of claim 1, which is encoded by a nucleic acid sequence which hybridizes under high stringency conditions with (i) nucleotides 214 to 2061 of SEQ ID NO. 1 or nucleotides 49 to 1944 of SEQ ID NO. 15, (ii) the genomic DNA sequence containing nucleotides 214 to 2061 of SEQ ID NO. 1 or nucleotides 49 to 1944 of SEQ ID NO. 15, (iii) a subsequence of (i) or (ii) of at least 100 nucleotides, or (iv) a complementary strand of (i), (ii), or (iii).

12. The polypeptide of claim 11, which is encoded by a nucleic acid sequence which hybridizes under high stringency conditions with (i) nucleotides 214 to 2061 of SEQ ID NO. 1 or nucleotides 49 to 1944 of SEQ ID NO. 15, (ii) the genomic DNA sequence containing nucleotides 214 to 2061 of SEQ ID NO. 1 or nucleotides 49 to 1944 of SEQ ID NO. 15, or (iii) a complementary strand of (i) or (ii).

21. A recombinant expression vector comprising the nucleic acid construct of claim 20.

22. A recombinant host cell comprising the nucleic acid construct of claim 20.

23. A method for producing a mutant nucleic acid sequence, comprising (a) introducing at least one mutation into the mature polypeptide coding sequence of SEQ ID NO. 1 or SEQ ID NO. 15, wherein the mutant nucleic acid sequence encodes a polypeptide consisting of amino acids 38 to 654 of SEQ ID NO. 2 or amino acids 17 to 648 of SEQ ID NO. 16, respectively; and (b) recovering the mutant nucleic acid sequence.

24. A mutant nucleic acid sequence produced by the method of claim 23.

25. A method for producing a polypeptide, comprising (a) cultivating a strain comprising the mutant nucleic acid sequence of claim 24 encoding the polypeptide to produce a supernatant comprising the polypeptide; and (b) recovering the polypeptide.

26. A method for producing the polypeptide of any of claims 1-16 comprising (a) cultivating a strain under conditions suitable for production of the polypeptide; and (b) recovering the polypeptide.

27. A method for producing the polypeptide of any of claims 1-16 comprising (a) cultivating a host cell comprising a nucleic acid construct comprising a nucleic acid sequence encoding the polypeptide under conditions suitable for production of the polypeptide; and (b) recovering the polypeptide.

28. A method for producing a polypeptide comprising (a) cultivating a host cell under conditions conducive for production of the polypeptide, wherein the host cell comprises a mutant nucleic acid sequence having at least one mutation in the mature polypeptide coding sequence of SEQ ID NO. 1 or SEQ ID NO. 15, wherein the mutant nucleic acid sequence encodes a polypeptide consisting of amino acids 38 to 654 of SEQ ID NO. 2 or

34.

36. A recombinant host cell comprising the nucleic acid construct of claim 34.

37. A method for producing a protein comprising (a) cultivating the recombinant host cell of claim 36 under conditions suitable for production of the protein; and (b) recovering the protein.

38. A method for reducing the phospholipid content in an edible oil, comprising treating the oil with the polypeptide of any of claims 1-16 to hydrolyze the phospholipid and separating the hydrolyzed phospholipid from the oil.

39. A method for reducing the phospholipid content of an aqueous carbohydrate solution or slurry, comprising treating the aqueous carbohydrate solution or slurry with the polypeptide of any of claims 1-16 to hydrolyze the phospholipid and separating the hydrolyzed phospholipid from the aqueous carbohydrate solution or slurry.

40. A method for preparing a dough, comprising incorporating into the dough an effective amount of the polypeptide of any of claims 1-16 to improve one or more properties of the dough and/or a baked product obtained from the dough..

41. The method of claim 40, wherein the one or more improved properties are selected from the group consisting of increased strength of the dough, increased stability of the dough, reduced stickiness of the dough, improved machinability of the dough, increased volume of the baked product, improved crumb structure of the baked product, improved softness of the baked product, improved flavor of the baked product, and improved antistaling of the baked product.

42. A method for preparing a baked product, comprising baking a dough produced by the method of claim 40 or 41 to produce a baked product.

43. A dough prepared by the method of claim 40 or 41.

```
TCTACACCTACTTGAATAGCTATATTCCCCGACATTTTATCAGCAATATTCAAGACATTCATATCAATACTTAGACATTG 80

GCTTGTTTCAACAAGCAACATGCTTGGCCCTCTCGTCTTTACTTTATGGCTTACCAGCTCGGCCATTGCTGCCCCGGATG 160
                     M  L  G  P  L  V  F  T  L  W  L  T  S  S  A  I  A  A  P  D

ACGCGGGTTTGGTCGCAGCACCAGCAATTGGCAAATCCCTCAGTATCCGGGCTCTTCCGGACTCTCCAAGCGGCGGTTAT 240
D  A  G  L  V  A  A  P  A  I  G  K  S  L  S  I  R  A  L  P  D  S  P  S  G  G  Y

GCGCCAAAAGTGGTTGACTGTCCCTCGACGCGCCCGAAAATCCGACTTGCCGATGGACTTTCAGACCAGGAAGAAGCCTG 320
 A  P  K  V  V  D  C  P  S  T  R  P  K  I  R  L  A  D  G  L  S  D  Q  E  E  A  W

GGTTCGTCGCCGAAGAAACAACACAATAGATCCAATGAAAGACTTGTTATCCCGAGTCAACATCTCGGGTTTCGACGCCG 400
 V  R  R  R  R  N  N  T  I  D  P  M  K  D  L  L  S  R  V  N  I  S  G  F  D  A

AAAAGTGGATTGAGAAAAACAAAAACAATGCGACTGCGCTACCTAACATCGCCATCGCAGCTTCTGGTGGTGGATACCGA 480
E  K  W  I  E  K  N  K  N  N  A  T  A  L  P  N  I  A  I  A  A  S  G  G  G  Y  R

GCCCTCATGAACGGAGCAGGCTTCATCTCTGCGGCTGACTCACGCAACAACGAATCCGGTCCCATCAGTGGTCTTCTACA 560
 A  L  M  N  G  A  G  F  I  S  A  A  D  S  R  N  N  E  S  G  P  I  S  G  L  L  Q

ATCTTCCACTTATTTGGCTGGTCTGTCAGGAGGTGGTTGGCTTGTTGGGTCTATCTTTGCCAACAACTTCACCACAATCC 640
 S  S  T  Y  L  A  G  L  S  G  G  G  W  L  V  G  S  I  F  A  N  N  F  T  T  I

CCGACCTACAAAAGGGAGACAAGGGTTCAGATATCTGGGCTTTTGACCGTTCAATTTTCAAAGGACCCGAAAAGTCAGGC 720
P  D  L  Q  K  G  D  K  G  S  D  I  W  A  F  D  R  S  I  F  K  G  P  E  K  S  G

TTGAACGTTTTGAACACGGCTAAATACTGGGATGACATAAAAGACACCGTTGACGAAAAGGCCGATGGGTGGAATACTAC 800
 L  N  V  L  N  T  A  K  Y  W  D  D  I  K  D  T  V  D  E  K  A  D  G  W  N  T  T
```

Fig. 1A

```
ACTCACTGACTGGTGGGGTCGTGCTCTGTCTTACCAGCTGATCGATGCCTCTGAGGGTGGTCCTGCGTATACTTTCTCCT 880
 L  T  D  W  W  G  R  A  L  S  Y  Q  L  I  D  A  S  E  G  G  P  A  Y  T  F  S

CCATCGCCGATACTTCCAACTTCAAGGACGCCGATACCCCATTTCCTATCCTCGTTGCCGATGGTCGTGCTCCTGGTCAA 960
S  I  A  D  T  S  N  F  K  D  A  D  T  P  F  P  I  L  V  A  D  G  R  A  P  G  Q

CGCATCGTTTCACTCAACGCAACTGTGTACGAGTTCAACCCGTTCGAGTTCGGAACATGGGATCCCACTAGTTACGGGTT 1040
 R  I  V  S  L  N  A  T  V  Y  E  F  N  P  F  E  F  G  T  W  D  P  T  S  Y  G  F

TGCTCCCGTCGAGTACATCGGCTCTAACTTCACGAATGGAACTATCGAAAAGGGCGGCGAATGTGTACGTGGCTTTGATC 1120
  A  P  V  E  Y  I  G  S  N  F  T  N  G  T  I  E  K  G  G  E  C  V  R  G  F  D

AGTTCGGCTTTGTTATGGGCACATCCTCCTCGTTATTCAACCAGTTCCTGCTCAACAACATCACCAAGATTGGAGAAGAA 1200
Q  F  G  F  V  M  G  T  S  S  S  L  F  N  Q  F  L  L  N  N  I  T  K  I  G  E  E

AATGACATTCCTTCACTTGTCGTGAAGGCTATCCAGGGATTCCTGGTAGCTTTGGATACCAATGATGAGGACATTGCGGA 1280
 N  D  I  P  S  L  V  V  K  A  I  Q  G  F  L  V  A  L  D  T  N  D  E  D  I  A  D

TTATTCTCCCAACCCATTCTACCAGTGGAACGTGACAGGAAAAAGCTACAACGCCAAGGACCATCAATTGACTCTTGTCG 1360
  Y  S  P  N  P  F  Y  Q  W  N  V  T  G  K  S  Y  N  A  K  D  H  Q  L  T  L  V

ACGGAGGGGAGGATCTGCAGAATATCCCACTCCATCCCTTGATCCAGCCTGTTCGTGGTGTCGACATCATCTTTGCCATC 1440
D  G  G  E  D  L  Q  N  I  P  L  H  P  L  I  Q  P  V  R  G  V  D  I  I  F  A  I

GATTCTTCAGCGGATACGGACAACAATTGGCCCAATGGTACCGCTCTTCGTGCGACATACGATCGTGTCGATTCCAGCTT 1520
 D  S  S  A  D  T  D  N  N  W  P  N  G  T  A  L  R  A  T  Y  D  R  V  D  S  S  L

AGGAAACGGAACTCAGTTTCCCTCTATTCCATCAGCTGAGACTTTCATTAATGAGAAGTTGAACCAACGCCCAACACTCT 1600
  G  N  G  T  Q  F  P  S  I  P  S  A  E  T  F  I  N  E  K  L  N  Q  R  P  T  L
```

Fig. 1B

```
TTGGCTGTGATGCAGACAACTTCACGCTTTCAGACGGCAAAGCTCCTCCTCCTCTTGTCTTCTACATTCCCAACGCGCCC 1680
 F  G  C  D  A  D  N  F  T  L  S  D  G  K  A  P  P  P  L  V  F  Y  I  P  N  A  P

TACACATTCTTGAGCAATGTCTCTACCTTCGATCTCTCATACAGCATCCCTGAGCGTGACAGTATCATTCTCAATGCTCT 1760
  Y  T  F  L  S  N  V  S  T  F  D  L  S  Y  S  I  P  E  R  D  S  I  I  L  N  A  L

GAACGGTGCCACTCAGGGCAATGGTACTATTGATAAGGAATGGCCCACGTGTGTCGTTTGTGCCATCATGAGCCGAAGTT 1840
   N  G  A  T  Q  G  N  G  T  I  D  K  E  W  P  T  C  V  V  C  A  I  M  S  R  S

GGTGGAAGTCCAATGAGACTGTTCCCAAAGAGTGCAGTACGTGTTTTGACAGATACTGCTGGGACGGAAAGTCGAATAAC 1920
W  W  K  S  N  E  T  V  P  K  E  C  S  T  C  F  D  R  Y  C  W  D  G  K  S  N  N

ACAGCTGTTAAGACTTACGAGCCTGGATTTATCATTGCTAACGAGACTGCGGCGGCGGAAGACAATGCAGCCGTTGCCGG 2000
  T  A  V  K  T  Y  E  P  G  F  I  I  A  N  E  T  A  A  A  E  D  N  A  A  V  A  G

TTTCACACCGAGTTGGTATATGTCTGTGTTTGTTGGAGTGGCTTTATTCTTGGCTATTTCGTAATTAGCATGTAGTTTCT 2080
   F  T  P  S  W  Y  M  S  V  F  V  G  V  A  L  F  L  A  I  S  .

GAGTTTGAATATATCTACATCATGCCAATAATATAATTAAAAAAAAAAAAAAAAAAAAAA 2141
```

Fig. 1C

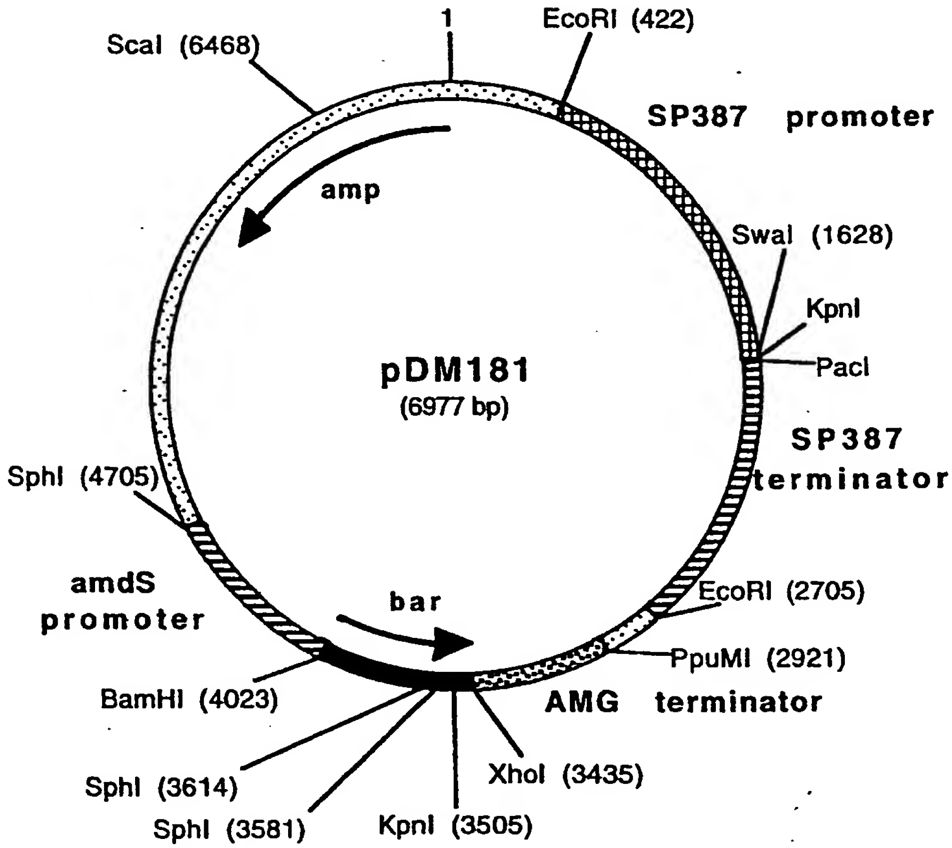
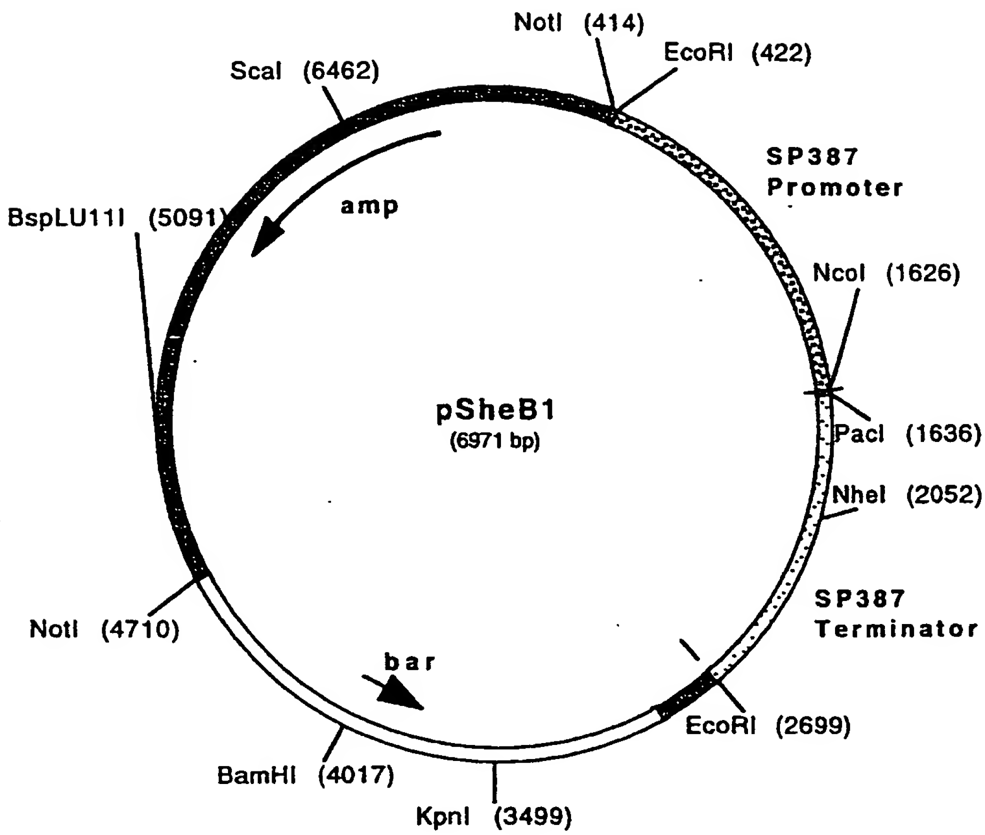

1
EcoRI (422)
ScaI (6468)
SP387 promoter
amp
SwaI (1628)
KpnI
PacI
pDM181
(6977 bp)
SP387
terminator
SphI (4705)
EcoRI (2705)
amdS
promoter
bar
PpuMI (2921)
BamHI (4023)
AMG terminator
XhoI (3435)
SphI (3614)
SphI (3581)
KpnI (3505)

Fig. 2



NotI (414)
EcoRI (422)
ScaI (6462)
SP387
Promoter
BspLU11I (5091)
amp
NcoI (1626)
pSheB1
(6971 bp)
PacI (1636)
NheI (2052)
SP387
Terminator
NotI (4710)
bar
EcoRI (2699)
BamHI (4017)
KpnI (3499)

Fig. 3

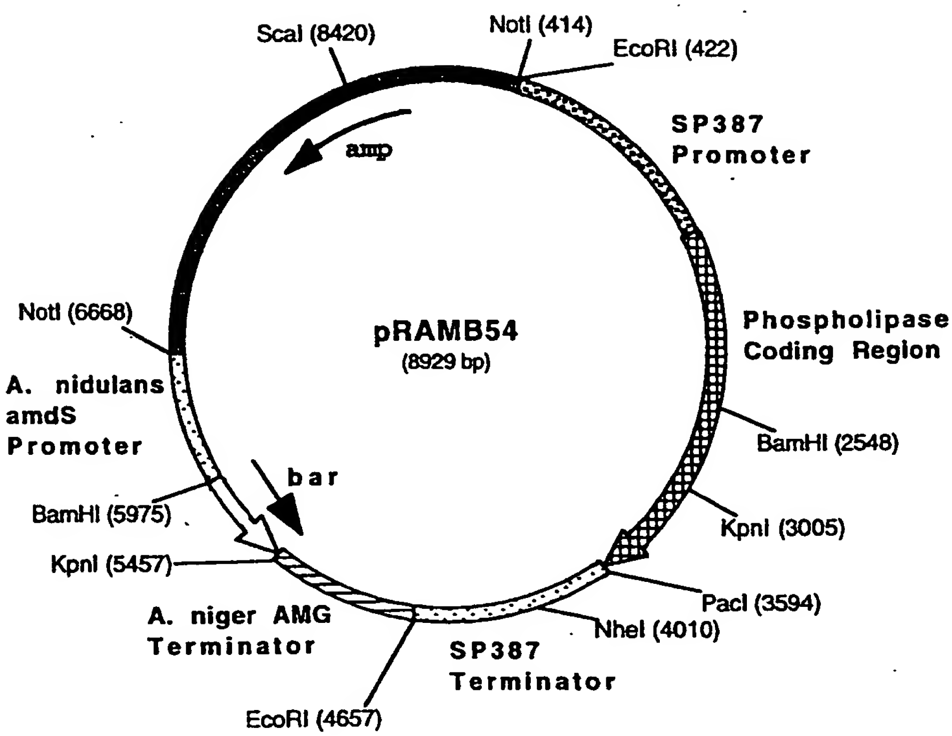

pRAMB54
(8929 bp)
ScaI (8420)
NotI (414)
EcoRI (422)
SP387 Promoter
amp
Phospholipase Coding Region
BamHI (2548)
KpnI (3005)
PacI (3594)
NheI (4010)
SP387 Terminator
EcoRI (4657)
A. niger AMG Terminator
KpnI (5457)
BamHI (5975)
bar
A. nidulans amdS Promoter
NotI (6668)

Fig. 4

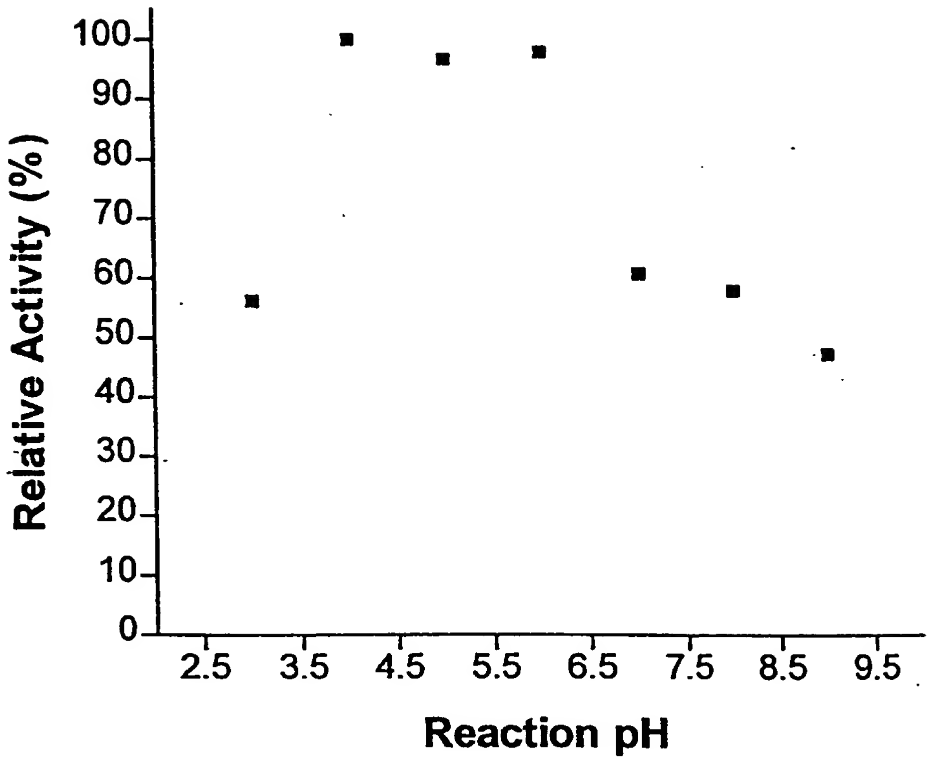

Relative Activity (%)
100
90
80
70
60
50
40
30
20
10
0
2.5
3.5
4.5
5.5
6.5
7.5
8.5
9.5
Reaction pH

Fig. 5

```
      M  L  G   F  V  A  L   T  L  W   L  S  T   A  I  A  A   P  E  D
  1   ATGCTCGGTT TTGTGGCCTT GACGCTCTGG CTCAGCACAG CCATTGCCGC TCCGGAAGAC
      T  A  L   I  P  R  V   N  S  V   E  I  R   A  L  P  N   S  P  S
 61   ACGGCTTTGA TCCCCAGGGT CAATTCAGTC GAGATCCGAG CCCTACCAAA TTCTCCGAGT
      G  G  Y   A  P  K  V   V  D  C   P  S  T   R  P  K  A   R  L  A
121   GGCGGATATG CTCCGAAAGT AGTCGATTGC CCTTCAACAC GTCCAAAAGC TCGTCTTGCC
      D  G  L   S  S  E  E   E  S  W   V  R  R   R  R  N  N   T  I  D
181   GATGGACTGT CTAGCGAGGA AGAATCCTGG GTTCGTCGTC GCAGAAACAA CACAATAGAT
      D  L  K   T  F  L  S   R  A  N   I  S  G   F  D  A  E   S  F  V
241   GACCTAAAGA CATTCTTGTC GCGAGCCAAC ATTTCCGGCT TCGATGCCGA GAGCTTCGTT
      E  K  H   K  N  N  A   T  G  L   P  N  I   A  I  A  A   S  G  G
301   GAGAAGCATA AAAACAATGC GACTGGTCTA CCCAACATCG CCATCGCCGC GTCCGGCGGT
      G  Y  R   A  L  M  N   G  A  G   F  L  S   A  A  D  S   R  N  N
361   GGTTATCGCG CTTTGATGAA CGGTGCAGGC TTTCTCTCAG CTGCTGATTC ACGCAACAAC
      K  T  G   P  I  S  G   L  L  Q   S  A  T   Y  L  A  G   L  S  G
421   AAGACCGGCC CCATCAGCGG CCTTCTCCAG TCAGCAACGT ATCTAGCTGG TCTTTCTGGC
      G  G  W   L  V  G  S   I  F  A   N  N  F   S  T  V  P   D  L  Q
481   GGTGGCTGGC TCGTTGGTTC TATTTTCGCC AACAATTTCT CAACAGTGCC AGACTTGCAA
      S  G  D   K  V  W  R   F  D  R   S  I  F   K  G  P  K   S  S  G
541   TCTGGTGACA AAGTCTGGCG ATTTGATCGC TCCATCTTCA AGGGCCCCAA GAGCTCTGGA
      I  S  L   L  N  T  A   E  Y  W   D  E  M   K  D  A  V   D  D  K
601   ATCAGTCTCC TCAACACTGC CGAGTACTGG GATGAGATGA AAGATGCCGT TGACGACAAG
      D  K  G   W  N  T  T   L  T  D   W  W  G   R  A  L  S   Y  Q  L
661   GACAAGGGCT GGAATACTAC TCTCACTGAT TGGTGGGGCC GCGCATTGTC GTATCAGCTC
      V  N  A   P  E  G  G   P  S  Y   T  F  S   S  I  A  D   T  S  N
721   GTCAATGCGC CTGAGGGTGG ACCTTCATAC ACTTTCTCTT CCATTGCCGA TACCTCCAAC
      F  K  D   A  D  T  P   F  P  I   L  V  A   D  G  R  A   P  G  E
781   TTCAAGGACG CAGACACACC CTTCCCCATC CTCGTTGCTG ATGGTCGTGC TCCAGGTGAG
```

Fig. 6A

```
        R  V  I   S  L  N  A   T  V  Y   E  F  N   P  Y  E  F   G  T  W
 841   CGCGTCATCT CCCTGAATGC TACTGTATAC GAGTTCAACC CATACGAGTT CGGTACATGG
        D  P  T   T  F  G  F   V  P  T   E  Y  L   A  S  N  F   T  N  G
 901   GATCCTACAA CCTTTGGCTT CGTACCCACT GAGTATCTCG CTTCCAACTT CACCAACGGC
        S  I  S   S  K  G  E   C  V  R   G  F  D   Q  I  G  F   V  M  G
 961   AGCATCAGCT CCAAGGGCGA ATGCGTTCGT GGCTTTGATC AGATCGGCTT TGTTATGGGA
        T  S  S   S  L  F  N   Q  F  L   L  N  N   I  T  K  V   G  K  E
1021   ACATCCTCTT CACTATTCAA CCAGTTCCTT CTCAACAACA TCACCAAGGT CGGCAAAGAG
        N  D  I   P  D  I  V   V  K  A   I  E  G   V  L  V  G   L  D  E
1081   AATGATATTC CCGACATTGT TGTCAAAGCC ATTGAAGGCG TCTTGGTTGG CTTGGACGAG
        D  D  E   D  I  A  Q   Y  A  P   N  P  F   F  G  W  N   P  T  D
1141   GATGATGAGG ATATCGCCCA GTATGCACCC AACCCCTTCT TCGGATGGAA CCCTACCGAT
       K  S  V  N  S  K  D   R  Q  L   T  L  V  D  G  G  E   D  L  Q
1201   AAGAGCGTCA ATTCCAAGGA CCGCCAGTTG ACTCTTGTCG ATGGTGGAGA GGATCTGCAG
        N  I  P   L  H  P  L   I  Q  P   V  R  G   V  D  I  I   F  A  I
1261   AACATTCCTC TTCACCCATT GATCCAGCCC GTTCGTGGCG TTGACATCAT TTTTGCCATT
        D  S  S   A  D  T  D   N  N  W   P  N  G   T  A  L  R   A  T  Y
1321   GATTCGTCTG CTGATACGGA CAACAACTGG CCCAATGGTA CTGCTCTTCG TGCTACATAT
        D  R  V   G  S  S  I   G  N  G   T  L  F   P  S  V  P   S  A  E
1381   GACCGTGTTG GTTCCAGCAT TGGCAACGGC ACACTATTTC CCTCAGTCCC ATCAGCTGAG
        T  F  I   N  E  K  L   N  Q  R   P  T  L   F  G  C  N   A  N  N
1441   ACCTTTATCA ACGAGAAGTT GAACCAGCGC CCTACGCTCT TTGGCTGTAA CGCGAACAAC
        F  T  L   S  D  G  E   V  P  P   P  L  I   L  Y  I  P   N  A  P
1501   TTCACTCTTT CAGACGGCGA AGTTCCCCCT CCGCTGATCT TGTACATCCC CAACGCTCCC
        Y  T  Y   H  S  N  V   S  T  F   D  M  S   Y  T  T  T   E  R  D
1561   TACACCTACC ACAGTAACGT CTCCACCTTC GACATGTCGT ATACCACTAC CGAACGCGAC
```

Fig. 6B

```
       N  I  I   L  N  A  L    N  G  A   T  Q  G   N  A  T  I    D  K  E
1621  AACATCATTC TTAATGCTCT CAATGGAGCT ACTCAGGGTA ACGCTACCAT TGACAAGGAG
       W  P  T   C  V  A  C    A  V  M   S  R  S   W  W  K  A    N  E  A
1681  TGGCCTACTT GCGTGGCGTG TGCTGTCATG AGCAGAAGCT GGTGGAAGGC GAATGAGGCT
       V  P  D   A  C  K  T    C  F  D   R  Y  C   W  D  G  K    S  N  N
1741  GTTCCTGACG CATGCAAGAC CTGCTTCGAT CGCTACTGCT GGGATGGAAA GTCTAACAAC
       T  A  V   K  S  Y  E    P  E  Y   I  I  G   G  N  A  T    A  E  A
1801  ACTGCTGTCA AGAGCTATGA GCCCGAGTAC ATCATTGGAG GAAATGCGAC CGCGGAAGCC
       A  D  N   A  A  G  A    R  L  G   P  S  W   F  V  S  A    G  V  G
1861  GCAGACAACG CTGCTGGAGC CAGGCTTGGT CCAAGCTGGT TTGTATCGGC TGGTGTCGGG
       A  A  A   L  F  A  L    M  *
1921  GCGGCTGCTT TGTTTGCTCT CATGTGA
```

Fig. 6C

SEQUENCE LISTING

<110> Randy M. Berka
Michael W. Rey
Tony Byun
Ryoko Itami
Noriko Tsutsumi
Alan Klotz

<120> Polypeptides Having Lysophospholipase Activity And Nucleic Acids Enocoding Same

<130> 5713.204-WO

<140> To Be Assigned
<141> 1999-11-10

<150> 09/189,486
<151> 1998-11-10

<160> 16

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 2141
<212> DNA
<213> Fusarium venenatum

<400> 1
```
tctacaccta cttgaatagc tatattcccc gacattttat cagcaatatt caagacattc     60
atatcaatac ttagacattg gcttgtttca acaagcaaca tgcttggccc tctcgtcttt    120
actttatggc ttaccagctc ggccattgct gccccggatg acgcgggttt ggtcgcagca    180
ccagcaattg gcaaatccct cagtatccgg gctcttccgg actctccaag cggcggttat    240
gcgccaaaag tggttgactg tccctcgacg cgcccgaaaa tccgacttgc cgatggactt    300
tcagaccagg aagaagcctg ggttcgtcgc cgaagaaaca acacaataga tccaatgaaa    360
gacttgttat cccgagtcaa catctcgggt ttcgacgccg aaaagtggat tgagaaaaac    420
aaaaacaatg cgactgcgct acctaacatc gccatcgcag cttctggtgg tggataccga    480
gccctcatga acggagcagg cttcatctct gcggctgact cacgcaacaa cgaatccggt    540
cccatcagtg gtcttctaca atcttccact tatttggctg gtctgtcagg aggtggttgg    600
cttgttgggt ctatctttgc caacaacttc accacaatcc ccgacctaca aaagggagac    660
aagggttcag atatctgggc ttttgaccgt tcaattttca aaggacccga aaagtcaggc    720
ttgaacgttt tgaacacggc taaatactgg gatgacataa aagacaccgt tgacgaaaag    780
gccgatgggt ggaatactac actcactgac tggtggggtc gtgctctgtc ttaccagctg    840
atcgatgcct ctgagggtgg tcctgcgtat actttctcct ccatcgccga tacttccaac    900
ttcaaggacg ccgatacccc atttcctatc ctcgttgccg atggtcgtgc tcctggtcaa    960
cgcatcgttt cactcaacgc aactgtgtac gagttcaacc cgttcgagtt cggaacatgg   1020
gatcccacta gttacgggtt tgctcccgtc gagtacatcg gctctaactt cacgaatgga   1080
actatcgaaa agggcggcga atgtgtacgt ggctttgatc agttcggctt tgttatgggc   1140
acatcctcct cgttattcaa ccagttcctg ctcaacaaca tcaccaagat tggagaagaa   1200
aatgacattc cttcacttgt cgtgaaggct atccagggat tcctggtagc tttggatacc   1260
aatgatgagg acattgcgga ttattctccc aacccattct accagtggaa cgtgacagga   1320
aaaagctaca acgccaagga ccatcaattg actcttgtcg acggagggga ggatctgcag   1380
```

aatatcccac tccatccctt gatccagcct gttcgtggtg tcgacatcat ctttgccatc 1440
gattcttcag cggatacgga caacaattgg cccaatggta ccgctcttcg tgcgacatac 1500
gatcgtgtcg attccagctt aggaaacgga actcagtttc cctctattcc atcagctgag 1560
actttcatta atgagaagtt gaaccaacgc ccaacactct ttggctgtga tgcagacaac 1620
ttcacgcttt cagacggcaa agctcctcct cctcttgtct tctacattcc caacgcgccc 1680
tacacattct tgagcaatgt ctctaccttc gatctctcat acagcatccc tgagcgtgac 1740
agtatcattc tcaatgctct gaacggtgcc actcagggca atggtactat tgataaggaa 1800
tggcccacgt gtgtcgtttg tgccatcatg agccgaagtt ggtggaagtc caatgagact 1860
gttcccaaag agtgcagtac gtgttttgac agatactgct gggacggaaa gtcgaataac 1920
acagctgtta agacttacga gcctggattt atcattgcta acgagactgc ggcggcggaa 1980
gacaatgcag ccgttgccgg tttcacaccg agttggtata tgtctgtgtt tgttggagtg 2040
gctttattct tggctatttc gtaattagca tgtagtttct gagtttgaat atatctacat 2100
catgccaata atataattaa aaaaaaaaaa aaaaaaaaaa a 2141

<210> 2
<211> 654
<212> PRT
<213> Fusarium venenatum

<400> 2
Met Leu Gly Pro Leu Val Phe Thr Leu Trp Leu Thr Ser Ser Ala Ile
1 5 10 15
Ala Ala Pro Asp Asp Ala Gly Leu Val Ala Ala Pro Ala Ile Gly Lys
20 25 30
Ser Leu Ser Ile Arg Ala Leu Pro Asp Ser Pro Ser Gly Gly Tyr Ala
35 40 45
Pro Lys Val Val Asp Cys Pro Ser Thr Arg Pro Lys Ile Arg Leu Ala
50 55 60
Asp Gly Leu Ser Asp Gln Glu Glu Ala Trp Val Arg Arg Arg Arg Asn
65 70 75 80
Asn Thr Ile Asp Pro Met Lys Asp Leu Leu Ser Arg Val Asn Ile Ser
85 90 95
Gly Phe Asp Ala Glu Lys Trp Ile Glu Lys Asn Lys Asn Asn Ala Thr
100 105 110
Ala Leu Pro Asn Ile Ala Ile Ala Ala Ser Gly Gly Gly Tyr Arg Ala
115 120 125
Leu Met Asn Gly Ala Gly Phe Ile Ser Ala Ala Asp Ser Arg Asn Asn
130 135 140
Glu Ser Gly Pro Ile Ser Gly Leu Leu Gln Ser Ser Thr Tyr Leu Ala
145 150 155 160
Gly Leu Ser Gly Gly Gly Trp Leu Val Gly Ser Ile Phe Ala Asn Asn
165 170 175
Phe Thr Thr Ile Pro Asp Leu Gln Lys Gly Asp Lys Gly Ser Asp Ile
180 185 190
Trp Ala Phe Asp Arg Ser Ile Phe Lys Gly Pro Glu Lys Ser Gly Leu
195 200 205
Asn Val Leu Asn Thr Ala Lys Tyr Trp Asp Asp Ile Lys Asp Thr Val
210 215 220
Asp Glu Lys Ala Asp Gly Trp Asn Thr Thr Leu Thr Asp Trp Trp Gly
225 230 235 240
Arg Ala Leu Ser Tyr Gln Leu Ile Asp Ala Ser Glu Gly Gly Pro Ala
245 250 255
Tyr Thr Phe Ser Ser Ile Ala Asp Thr Ser Asn Phe Lys Asp Ala Asp
260 265 270
Thr Pro Phe Pro Ile Leu Val Ala Asp Gly Arg Ala Pro Gly Gln Arg
275 280 285

Ile Val Ser Leu Asn Ala Thr Val Tyr Glu Phe Asn Pro Phe Glu Phe
290 295 300
Gly Thr Trp Asp Pro Thr Ser Tyr Gly Phe Ala Pro Val Glu Tyr Ile
305 310 315 320
Gly Ser Asn Phe Thr Asn Gly Thr Ile Glu Lys Gly Gly Glu Cys Val
325 330 335
Arg Gly Phe Asp Gln Phe Gly Phe Val Met Gly Thr Ser Ser Ser Leu
340 345 350
Phe Asn Gln Phe Leu Leu Asn Asn Ile Thr Lys Ile Gly Glu Glu Asn
355 360 365
Asp Ile Pro Ser Leu Val Val Lys Ala Ile Gln Gly Phe Leu Val Ala
370 375 380
Leu Asp Thr Asn Asp Glu Asp Ile Ala Asp Tyr Ser Pro Asn Pro Phe
385 390 395 400
Tyr Gln Trp Asn Val Thr Gly Lys Ser Tyr Asn Ala Lys Asp His Gln
405 410 415
Leu Thr Leu Val Asp Gly Gly Glu Asp Leu Gln Asn Ile Pro Leu His
420 425 430
Pro Leu Ile Gln Pro Val Arg Gly Val Asp Ile Ile Phe Ala Ile Asp
435 440 445
Ser Ser Ala Asp Thr Asp Asn Asn Trp Pro Asn Gly Thr Ala Leu Arg
450 455 460
Ala Thr Tyr Asp Arg Val Asp Ser Ser Leu Gly Asn Gly Thr Gln Phe
465 470 475 480
Pro Ser Ile Pro Ser Ala Glu Thr Phe Ile Asn Glu Lys Leu Asn Gln
485 490 495
Arg Pro Thr Leu Phe Gly Cys Asp Ala Asp Asn Phe Thr Leu Ser Asp
500 505 510
Gly Lys Ala Pro Pro Pro Leu Val Phe Tyr Ile Pro Asn Ala Pro Tyr
515 520 525
Thr Phe Leu Ser Asn Val Ser Thr Phe Asp Leu Ser Tyr Ser Ile Pro
530 535 540
Glu Arg Asp Ser Ile Ile Leu Asn Ala Leu Asn Gly Ala Thr Gln Gly
545 550 555 560
Asn Gly Thr Ile Asp Lys Glu Trp Pro Thr Cys Val Val Cys Ala Ile
565 570 575
Met Ser Arg Ser Trp Trp Lys Ser Asn Glu Thr Val Pro Lys Glu Cys
580 585 590
Ser Thr Cys Phe Asp Arg Tyr Cys Trp Asp Gly Lys Ser Asn Asn Thr
595 600 605
Ala Val Lys Thr Tyr Glu Pro Gly Phe Ile Ile Ala Asn Glu Thr Ala
610 615 620
Ala Ala Glu Asp Asn Ala Ala Val Ala Gly Phe Thr Pro Ser Trp Tyr
625 630 635 640
Met Ser Val Phe Val Gly Val Ala Leu Phe Leu Ala Ile Ser
645 650

<210> 3
<211> 30
<212> DNA
<213> Fusarium venenatum

<400> 3
gagctcgagg aattcttaca aaccttcaac 30

<210> 4

<211> 47
<212> DNA
<213> Fusarium venenatum

<400> 4
ttaattaagg tacctgaatt taaatggtga agagatagat atccaag 47

<210> 5
<211> 51
<212> DNA
<213> Fusarium venenatum

<400> 5
tcaccattta aattcaggta ccttaattaa attccttgtt ggaagcgtcg a 51

<210> 6
<211> 42
<212> DNA
<213> Fusarium venenatum

<400> 6
tggtatgcat aagcttgaat tcaggtaaac aagatataat tt 42

<210> 7
<211> 35
<212> DNA
<213> Fusarium venenatum

<400> 7
cagtgaattg gcctcgatgg ccgcggccgc gaatt 35

<210> 8
<211> 35
<212> DNA
<213> Fusarium venenatum

<400> 8
aattcgcggc cgcggccatc gaggccaatt cactg 35

<210> 9
<211> 34
<212> DNA
<213> Fusarium venenatum

<400> 9
cacgaaggaa agacgatggc tttcacggtg tctg 34

<210> 10
<211> 34
<212> DNA
<213> Fusarium venenatum

<400> 10
cagacaccgt gaaagccatc gtctttcctt cgtg 34

<210> 11

<211> 46
<212> DNA
<213> Fusarium venenatum

<400> 11
ctatctcttc accatggtac cttaattaaa taccttgttg gaagcg 46

<210> 12
<211> 46
<212> DNA
<213> Fusarium venenatum

<400> 12
cgcttccaac aaggtattta attaaggtac catggtgaag agatag 46

<210> 13
<211> 29
<212> DNA
<213> Fusarium venenatum

<400> 13
ggcacatgtt gggccctctc gtctttact 29

<210> 14
<211> 36
<212> DNA
<213> Fusarium venenatum

<400> 14
gacttaatta atttacgaaa tagccaagaa taaagc 36

<210> 15
<211> 1947
<212> DNA
<213> Fusarium verticillioides

<400> 15
atgctcggtt ttgtggcctt gacgctctgg ctcagcacag ccattgccgc tccggaagac 60
acggctttga tccccagggt caattcagtc gagatccgag ccctaccaaa ttctccgagt 120
ggcggatatg ctccgaaagt agtcgattgc ccttcaacac gtccaaaagc tcgtcttgcc 180
gatggactgt ctagcgagga agaatcctgg gttcgtcgtc gcagaaacaa cacaatagat 240
gacctaaaga cattcttgtc gcgagccaac atttccggct tcgatgccga gagcttcgtt 300
gagaagcata aaaacaatgc gactggtcta cccaacatcg ccatcgccgc gtccggcggt 360
ggttatcgcg ctttgatgaa cggtgcaggc tttctctcag ctgctgattc acgcaacaac 420
aagaccggcc ccatcagcgg ccttctccag tcagcaacgt atctagctgg tctttctggc 480
ggtggctggc tcgttggttc tattttcgcc aacaatttct caacagtgcc agacttgcaa 540
tctggtgaca aagtctggcg atttgatcgc tccatcttca agggccccaa gagctctgga 600
atcagtctcc tcaacactgc cgagtactgg gatgagatga aagatgccgt tgacgacaag 660
gacaagggct ggaatactac tctcactgat tggtggggcc gcgcattgtc gtatcagctc 720
gtcaatgcgc ctgagggtgg accttcatac actttctctt ccattgccga tacctccaac 780
ttcaaggacg cagacacacc cttccccatc ctcgttgctg atggtcgtgc tccaggtgag 840
cgcgtcatct ccctgaatgc tactgtatac gagttcaacc catacgagtt cggtacatgg 900
gatcctacaa cctttggctt cgtacccact gagtatctcg cttccaactt caccaacggc 960
agcatcagct ccaagggcga atgcgttcgt ggctttgatc agatcggctt tgttatggga 1020
acatcctctt cactattcaa ccagttcctt ctcaacaaca tcaccaaggt cggcaaagag 1080
aatgatattc ccgacattgt tgtcaaagcc attgaaggcg tcttggttgg cttggacgag 1140

```
gatgatgagg atatcgccca gtatgcaccc aacccccttct tcggatggaa ccctaccgat     1200
aagagcgtca attccaagga ccgccagttg actcttgtcg atggtggaga ggatctgcag     1260
aacattcctc ttcacccatt gatccagccc gttcgtggcg ttgacatcat ttttgccatt     1320
gattcgtctg ctgatacgga caacaactgg cccaatggta ctgctcttcg tgctacatat     1380
gaccgtgttg gttccagcat tggcaacggc acactatttc cctcagtccc atcagctgag     1440
acctttatca acgagaagtt gaaccagcgc cctacgctct ttggctgtaa cgcgaacaac     1500
ttcactcttt cagacggcga agttcccccct ccgctgatct tgtacatccc caacgctccc     1560
tacacctacc acagtaacgt ctccaccttc gacatgtcgt ataccactac cgaacgcgac     1620
aacatcattc ttaatgctct caatggagct actcagggta acgctaccat tgacaaggag     1680
tggcctactt gcgtggcgtg tgctgtcatg agcagaagct ggtggaaggc gaatgaggct     1740
gttcctgacg catgcaagac ctgcttcgat cgctactgct gggatggaaa gtctaacaac     1800
actgctgtca agagctatga gcccgagtac atcattggag gaaatgcgac cgcggaagcc     1860
gcagacaacg ctgctggagc caggcttggt ccaagctggt ttgtatcggc tggtgtcggg     1920
gcggctgctt tgtttgctct catgtga                                          1947
```

```
<210> 16
<211> 648
<212> PRT
<213> Fusarium verticillioides

<400> 16
Met Leu Gly Phe Val Ala Leu Thr Leu Trp Leu Ser Thr Ala Ile Ala
1               5                   10                  15
Ala Pro Glu Asp Thr Ala Leu Ile Pro Arg Val Asn Ser Val Glu Ile
            20                  25                  30
Arg Ala Leu Pro Asn Ser Pro Ser Gly Gly Tyr Ala Pro Lys Val Val
        35                  40                  45
Asp Cys Pro Ser Thr Arg Pro Lys Ala Arg Leu Ala Asp Gly Leu Ser
    50                  55                  60
Ser Glu Glu Glu Ser Trp Val Arg Arg Arg Arg Asn Asn Thr Ile Asp
65                  70                  75                  80
Asp Leu Lys Thr Phe Leu Ser Arg Ala Asn Ile Ser Gly Phe Asp Ala
                85                  90                  95
Glu Ser Phe Val Glu Lys His Lys Asn Asn Ala Thr Gly Leu Pro Asn
            100                 105                 110
Ile Ala Ile Ala Ala Ser Gly Gly Gly Tyr Arg Ala Leu Met Asn Gly
        115                 120                 125
Ala Gly Phe Leu Ser Ala Ala Asp Ser Arg Asn Asn Lys Thr Gly Pro
    130                 135                 140
Ile Ser Gly Leu Leu Gln Ser Ala Thr Tyr Leu Ala Gly Leu Ser Gly
145                 150                 155                 160
Gly Gly Trp Leu Val Gly Ser Ile Phe Ala Asn Asn Phe Ser Thr Val
                165                 170                 175
Pro Asp Leu Gln Ser Gly Asp Lys Val Trp Arg Phe Asp Arg Ser Ile
            180                 185                 190
Phe Lys Gly Pro Lys Ser Ser Gly Ile Ser Leu Leu Asn Thr Ala Glu
        195                 200                 205
Tyr Trp Asp Glu Met Lys Asp Ala Val Asp Asp Lys Asp Lys Gly Trp
    210                 215                 220
Asn Thr Thr Leu Thr Asp Trp Trp Gly Arg Ala Leu Ser Tyr Gln Leu
225                 230                 235                 240
Val Asn Ala Pro Glu Gly Gly Pro Ser Tyr Thr Phe Ser Ser Ile Ala
                245                 250                 255
Asp Thr Ser Asn Phe Lys Asp Ala Asp Thr Pro Phe Pro Ile Leu Val
            260                 265                 270
Ala Asp Gly Arg Ala Pro Gly Glu Arg Val Ile Ser Leu Asn Ala Thr
```

```
        275                 280                 285
Val Tyr Glu Phe Asn Pro Tyr Glu Phe Gly Thr Trp Asp Pro Thr Thr
    290                 295                 300
Phe Gly Phe Val Pro Thr Glu Tyr Leu Ala Ser Asn Phe Thr Asn Gly
305                 310                 315                 320
Ser Ile Ser Ser Lys Gly Glu Cys Val Arg Gly Phe Asp Gln Ile Gly
                325                 330                 335
Phe Val Met Gly Thr Ser Ser Ser Leu Phe Asn Gln Phe Leu Leu Asn
            340                 345                 350
Asn Ile Thr Lys Val Gly Lys Glu Asn Asp Ile Pro Asp Ile Val Val
        355                 360                 365
Lys Ala Ile Glu Gly Val Leu Val Gly Leu Asp Glu Asp Asp Glu Asp
    370                 375                 380
Ile Ala Gln Tyr Ala Pro Asn Pro Phe Phe Gly Trp Asn Pro Thr Asp
385                 390                 395                 400
Lys Ser Val Asn Ser Lys Asp Arg Gln Leu Thr Leu Val Asp Gly Gly
                405                 410                 415
Glu Asp Leu Gln Asn Ile Pro Leu His Pro Leu Ile Gln Pro Val Arg
            420                 425                 430
Gly Val Asp Ile Ile Phe Ala Ile Asp Ser Ser Ala Asp Thr Asp Asn
        435                 440                 445
Asn Trp Pro Asn Gly Thr Ala Leu Arg Ala Thr Tyr Asp Arg Val Gly
    450                 455                 460
Ser Ser Ile Gly Asn Gly Thr Leu Phe Pro Ser Val Pro Ser Ala Glu
465                 470                 475                 480
Thr Phe Ile Asn Glu Lys Leu Asn Gln Arg Pro Thr Leu Phe Gly Cys
                485                 490                 495
Asn Ala Asn Asn Phe Thr Leu Ser Asp Gly Glu Val Pro Pro Pro Leu
            500                 505                 510
Ile Leu Tyr Ile Pro Asn Ala Pro Tyr Thr Tyr His Ser Asn Val Ser
        515                 520                 525
Thr Phe Asp Met Ser Tyr Thr Thr Thr Glu Arg Asp Asn Ile Ile Leu
    530                 535                 540
Asn Ala Leu Asn Gly Ala Thr Gln Gly Asn Ala Thr Ile Asp Lys Glu
545                 550                 555                 560
Trp Pro Thr Cys Val Ala Cys Ala Val Met Ser Arg Ser Trp Trp Lys
                565                 570                 575
Ala Asn Glu Ala Val Pro Asp Ala Cys Lys Thr Cys Phe Asp Arg Tyr
            580                 585                 590
Cys Trp Asp Gly Lys Ser Asn Asn Thr Ala Val Lys Ser Tyr Glu Pro
        595                 600                 605
Glu Tyr Ile Ile Gly Gly Asn Ala Thr Ala Glu Ala Ala Asp Asn Ala
    610                 615                 620
Ala Gly Ala Arg Leu Gly Pro Ser Trp Phe Val Ser Ala Gly Val Gly
625                 630                 635                 640
Ala Ala Ala Leu Phe Ala Leu Met
                645
```

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 99/26789

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/55 C12N9/20 A23D7/00 A23L1/10

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N A23L A23D

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 98 26057 A (NOVONORDISK AS) 18 June 1998 (1998-06-18) the whole document	1-44
A	CATCHESIDE D.E.A ET AL.: " Neurospora crassa Lindegren 25a lysophospholipase (lpl) gene" EMBL DATABSE ; ACCESSION NUMBER AF045575, 20 February 1998 (1998-02-20), XP002135649 the whole document	1-22

☐ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
13 April 2000	28/04/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Gurdjian, D

4

INTERNATIONAL SEARCH REPORT

International Application No. PCT/ US 99/26789

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-44
lysophospholipase from Fusarium venetatum, its cloned DNA sequence encoding the enzyme a method of producing the enzyme, and the use of said enzyme for a number of industrial applications .

2. Claims: 1-44
lysophospholipase from Fusarium verticillioides, its cloned DNA sequence encoding the enzyme a method of producing the enzyme, and the use of said enzyme for a number of industrial applications .

PATENT COOPERATION TREA1

From the INTERNATIONAL SEARCHING AUTHORITY

PCT

To:

NOVO NORDISK OF NORTH AMERICA, INC.
Attn. ZELSON, Steve
405 Lexington Avenue
Suite 6400
New York, New York 10174
UNITED STATES OF AMERICA

RECEIVED FEB 29 2000

INVITATION TO FURNISH NUCLEOTIDE AND/OR AMINO ACID SEQUENCE LISTING COMPLYING WITH WIPO STANDARD ST25

(PCT Rule 13ter.1(a) and (c) and Administrative Instructions, Section 208 and Annex C)

	Date of mailing *(day/month/year)* 23/02/2000
Applicant's or agent's file reference 5713.204-WO	REPLY DUE within 1 months/days from the above date of mailing
International application No. PCT/US 99/26789	International filing date *(day/month/year)* 10/11/1999
Applicant NOVO NORDISK BIOTECH, INC. et al.	

1. The applicant is hereby invited, within the time limit indicated above, to furnish to this Authority:

[ ] a nucleotide and/or amino acid sequence listing in written form complying with the standard provided for in Annex C of the Administrative Instructions, accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.

[X] a statement to the effect that the sequence listing in written form, already furnished to this Authority, does not go beyond the disclosure in the international application as filed.

[X] a nucleotide and/or amino acid sequence listing in computer readable form complying with the standard provided for in Annex C of the Administrative Instructions, accompanied by a statement that the information recorded in computer readable form is identical to the written sequence listing.

[ ] a statement that the information recorded in computer readable form (that computer readable form having already been furnished to this Authority) is identical to the written sequence listing.

2. Failure to comply with this invitation may result in this Authority not carrying out the international search to the extent that no meaningful search can be carried out.

3. Further observations (if necessary):

IMPORTANT REMARK
The statements are legally required [See Suppl No 2 to Official Journal No 11/1998 (page 14, ¶ 37 & 40 and page 64 ¶ III.2)]

Name and mailing address of the International Searching Authority European Patent Office, P.B. 5818 Patentlaan 2 NL-2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Authorized officer Mirella Deleye-Milani M. Deleye

Form PCT/ISA/225 (July 1998)

NZAS-0027377

Office Européen des Brevets DG1
Europäisches PatentAmt DG1
European Patent Office DG1

Storage and Retrieval of Amino acid and Nucleotide Data

P.B. 5818 Patentlaan 2 NL-2280 HV Rijswijk
Fax : + 31 70 340 39 92

ANNEX

Dear applicant/representative,

According to Supplement 2 to the Official Journal Nr.11/98 of the EPO (& Rules 13 ter & 5.2. PCT), if nucleotide/amino acid sequences are disclosed in a European/International patent application, the description shall contain a sequence listing complying with WIPO ST.25.

Thank you for the filed document on paper form.

However, since now, we didn't received the computer readable form.

In order not to delay the search anylonger, we herewith kindy invite you to submit a **sequence listing in computer readable form accompanied by the appropriate statements.**

Relating to this, we remind you that if these requirements are not met or not met in due time, the EPO does not perform the international search where a meaningful search cannot be carried out (Rule 13ter.1(c)PCT). In this case the international search report is replaced in full or in part by the statement under Article 17(2)(a)(ii)PCT.

The computer readable form of the Sequence Listing in ASCII format (text only) is mandatory.

For further questions do not hesitate to contact us.

Please send sequence listing on paper and/or in computer-readable form preferably to the European Patent Office, STRAND Program, Directorate Biotechnology (Dir 1212), Attn. : Ms M. Deleye, Room S 02 N 24 - Patentlaan 2, NL 2288 EE RIJSWIJK (NL)